

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
 RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
 THE PHILIPS INDUSTRIES

APPLIED STATISTICS

AN IMPORTANT PHASE IN THE DEVELOPMENT OF EXPERIMENTAL SCIENCE

by H. C. HAMAKER.

31:519.2

*On 20th May 1960 Dr. H. C. Hamaker presented an address under the above title to mark his inauguration as extra-mural professor at the Technische Hogeschool, Eindhoven. As we have frequently done on similar occasions in the past, we print below the text of his discourse in extenso *). The author has introduced some changes here and there to suit a wider audience than that which attended the inaugural ceremony. Some years ago a series of articles on sampling systems from the pen of Prof. Hamaker were published in this journal **). These articles were largely confined to lot inspection. In recent years statistical methods have been more and more applied to research in general: the author traces here the broad lines of this development. A short bibliography is appended of articles on statistics contributed by the author in recent years to professional journals.*

Fifteen or so years ago the statistician in industry was still virtually unknown in this country and in many others. The notion that statistics might be usefully applied to technological problems was more often than not, as I myself discovered, dismissed with a shrug of the shoulders.

Since then, however, the situation has radically changed. The Philips Research Laboratories at Eindhoven now employ a group of eight or ten professional statisticians who are engaged solely on research in this field and on the application of statistical ideas and methods to industrial problems. Moreover, scattered throughout the company there are smaller groups of statisticians working on problems specific to a particular division and its products. Similar developments have taken place in other industries at home and abroad, and the market has not yet by any means reached saturation point. There is a constant demand for experienced statisticians, and for training and instruction at various levels that will enable industrial personnel to solve simple problems themselves and to judge when the services of a professional statistician can best be called upon. This is not a passing vogue.

On the contrary, I believe that we are in the middle of an important phase in the evolution of scientific thought, which will have a lasting influence on almost every branch of science. I should like first of all to dwell on this aspect for a moment.

Technology, the arbiter and pacemaker of our modern society, is a product of the rapid advances made in the natural sciences. These in their turn sprang from man's innate urge to understand and explain the phenomena he perceives in the world around him. The first steps in this direction are to make objective observations, and to record these observations in terms of numerical quantities which can be compared with other similar quantities and used to perform mathematical operations.

Experience has shown, however, that numerical observations are hardly ever exactly reproducible; when an experiment is repeated the results invariably show a certain spread. Accurate results are obtained only when the observations are made under carefully chosen, controlled conditions. The scientist would therefore retire behind the walls of special laboratories where such conditions could be established. The utmost care was devoted to the construction of the equipment employed, and the objects to be studied were carefully selected; only those were admitted that appeared to allow accurate and reproducible observations. Although no observation is ever perfectly reproducible, no matter how elabo-

*) Published by Centrex, Eindhoven, 1960.

**) H. C. Hamaker, Lot inspection by sampling; The theory of sampling inspection plans; The practical application of sampling inspection plans and tables (with J. J. M. Taudin Chabot and F. G. Willemze), Philips tech. Rev. 11, 176-182, 260-270, 362-370, 1949/50.

rate the precautions taken, the accuracy achieved was such that the results could be interpreted without taking their spread into account. To facilitate the interpretation of the results the problems chosen were simple, involving only a limited number of factors. "Exact", reproducible observation, with only one factor varied at a time, became the ideal that typified the classic exact sciences. Hence the qualification *exact*.

It is in the nature of man to tend to overestimate the significance of methods and ways of thinking that have been successful, and to attribute to them a general validity beyond the confines of the field for which they were devised. The success of the exact sciences accordingly led to the belief that the path followed was the only true one, and those who followed it were inclined to look down on all investigations that had to be content with less accurate observations.

In this they were wrong, for there are many fields of research where exact observation is essentially impossible; for example, the branches of science concerned with living beings. We can, it is true, determine with some accuracy the metabolism of a single plant or animal, but the value of such observations depends on the extent to which the results may be interpreted as valid for the species as a whole. And such interpretations must inevitably take account of the variations from one individual to another, which are always considerable.

Again, it is often wrong to limit experiments to the variation of only one factor. Agricultural experiments are a case in point. Suppose we want to find out how the yield of different varieties of potatoes depends on the nature and quantity of manure. We must not investigate separately the influence of manure containing potassium and of manure containing nitrogen, because their combined effect cannot be predicted from their separate effects, i.e. there is an "interaction" between these variables. Nor can the experiment be divided into parts to be done in different years, for changeable weather conditions will make it impossible to compare the results. The only efficient experimental method is to deal with all the varieties and manures in one season. The elements of this experiment should be distributed over a limited area of land in such a way as to minimize unavoidable variations in the fertility of the soil. Even then, fairly important random fluctuations will continue to influence the results, and these must be taken into account by an appropriate mathematical method of interpretation.

Between 1920 and 1930, the well-known English statistician Sir Ronald Fisher, at the Agricultural

Experimental Station at Rothamstead, was the first to recognize that the principles of agricultural experimentation should be entirely different from those of the "exact" experiments in the laboratory. In his book "The design of experiments" he showed the way which is still being pursued¹⁾.

According to these new principles there is nothing against varying several factors simultaneously in a single experiment, provided the layout is so designed that the effect of each factor can be separately established by an appropriate analysis of the observations. A further merit of such designs is that the extent to which the various factors influence one another can also be determined. This is of course impossible if only one factor is varied.

The random fluctuations impart an element of uncertainty to any conclusion drawn from the observations. This implies accepting that there is always a certain risk of drawing a wrong conclusion. This is no objection provided the design of the experiment permits an analysis of the magnitude of that risk: we are then in a position to keep it within reasonable bounds.

In these experimental designs statistical techniques of analysis and interpretation play a most important role. They are derived from the statistical principles evolved at the turn of the century by Galton and Pearson, which are based mainly on the theory of probability²⁾. This explains why the new ideas have been developed by statisticians and are nowadays principally described in text books on statistics.

In my view, however, they should really be seen as a fundamental change in the principles of experimental science. The classic ideals of exactitude, and of varying only one factor at a time, forced the experimenter into a straitjacket, from which he was freed by the work of Fisher. Phenomena formerly excluded from the preserves of exact science, for being inaccessible to exact and reproducible observation, can now be subjected to a mathematically well-founded — and thus in a certain sense "exact" — treatment. This has considerably widened and deepened the field of applied science. No wonder, then, that statistics nowadays plays an active part in almost every branch of knowledge.

Although in the application of statistics the line of thought remains consistent, the methods adopted

¹⁾ R. A. Fisher, *The design of experiments*, Oliver and Boyd, Edinburgh 1935.

²⁾ See e.g. K. Pearson, *The life, letters and labours of Francis Galton*, Cambridge Univ. Press 1914 (Part I), 1924 (Part II), 1930 (Part IIIa, b).

E. S. Pearson, *Karl Pearson, an appreciation of some aspects of his life and work*, Cambridge Univ. Press 1938.

are many and various and must always be adapted to the given problems and circumstances. To take the case of *agriculture* again, only one experiment a year can be carried out; elaborate and complicated experimental designs are therefore justified. The experiment itself is done in spring and summer, the observations are harvested in the autumn, and the winter is spent in analysing them and in designing a fresh experiment.

In the *electrical engineering industry*, on the other hand, we are concerned with machines capable of turning out a thousand and more products an hour. There is no need here to skimp the number of observations, but the experiments must be kept as simple as possible to avoid organizational difficulties and laborious analysis. At such a rate of production the engineer is in a hurry, and he would really have preferred the answer yesterday to the experiments being done today.

Between these extremes we have the *chemical industry*, which may process one charge a day. Here the investigator must be sparing in his observations, because they are costly and because a comprehensive investigation would take too long.

In *medical research* we may want to compare two medicines: neither the patient nor his doctor must know which one has been administered, to preclude the possibility that their judgement of the results is coloured by preconceived opinions. Nor is it possible to carry out the experiment as and when it suits the investigator; its rate of progress is determined by the availability of patients. This has the advantage, however, that the results can be gradually accumulated and the experiment stopped as soon as a clear conclusion has been reached.

Again, in *psychological experiments* we are not primarily concerned with studying the effects of certain factors. The problems involved are entirely different. There are, for example, many possible methods of estimating human intelligence, but intelligence cannot be *measured*; it is only by comparing the various methods one with the other that we can try to discover in how far they answer their purpose.

The *sociologist* has to rely on surveys for his observations; statistical methods are important here for processing and interpreting the results, and also for devising an efficient survey technique.

I could go on like this for some time, but I shall turn now from the general to the particular, in order to illustrate some of the ideas and methods which form the statistician's stock-in-trade. I shall take the example that first springs to mind on this occasion,

namely the address I am now delivering and inaugural addresses in general.

Let us imagine we want to know, for one reason or another, the average length of the words used in this discourse, length meaning the number of letters per word. How can we best find it?

My lecture consists of a collection of 6000 to 7000 words, and no-one will expect me to have measured them all. It is evidently possible to determine the average word-length with a fair accuracy by counting a limited number of words. We start by counting 200 words, picked at random from the text; in the terminology of statistics we take a "sample" of 200 words. The result is set forth in fig. 1 in a

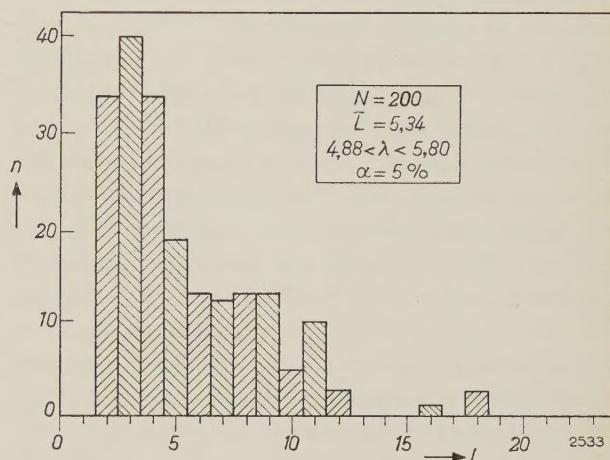


Fig. 1. Histogram of counted word-lengths in a random sample (drawn by lot) of 200 words. n = number of times the word-length L (= number of letters per word) is found.

so-called histogram. Of the 200 words, 34 were found to contain 2 letters, 40 had 3 letters, 34 had 4, and so on; the longest word, with 18 letters, was observed three times *). These 200 observations yield an average length

$$\bar{L} = 5.34 \text{ letters per word.}$$

Of course, this value does not agree exactly with the "true" average length λ which we should find if we counted all the words. Nevertheless, from the dispersion of the results, as appears in the figure, we can establish limits between which this unknown parameter (λ) can be said to lie with a given degree of confidence. This yields:

$$4.88 < \lambda < 5.80 \text{ letters per word}$$

with a confidence of 95%.

If I assert that λ lies between these limits, I run a 5% risk of telling an untruth. The limits form a so-called *confidence interval*. For the confidence level we can choose any value we like; a higher level of, say, 99% will lead to a wider interval. The width of

*) The analysis refers to the [original Dutch version of this lecture.]

the interval is inversely proportional to the root of the number of observations, so that we must make four times as many observations if we wish to halve the width of the interval; i.e. 800 instead of 200.

Two important questions now arise. Firstly, how are we to pick out the words that will constitute the sample? And secondly, what degree of accuracy, or what width of confidence interval, should we aim at; in other words, how big must the sample be?

To the first question the statistician has a ready answer. The sample must be a random one; it may be obtained by numbering the words of this speech consecutively from beginning to end and determining the serial numbers of the words that are to make up the sample by a lottery.

There is nothing difficult about that, for every applied statistician worth his salt carries a lottery about with him. A simple example of such a device consists of a glass bulb with a hollow stem and filled with ten wooden beads, nine of them green and one red. When the beads are shaken from the bulb into the stem, the position of the red bead indicates a number between 0 and 9. By repeating this operation many times we obtain a set of *random numbers*, that is numbers from 0 to 9 in completely random order. Groups of four random numbers together then indicate the serial numbers of the words to be included in the sample; numbers higher than the total of words in the text are ignored.

Plainly, this method of sampling is extremely cumbersome, and in the present case unnecessary. In any language, long and short words succeed one another with a fair degree of regularity. Perhaps, in an inaugural address of this nature, one might expect to find rather more long words in the middle, when the speaker has warmed up to his specific technical jargon, than at the beginning or end, but this can have no pronounced effect. If, to be on the safe side, we choose one or two lines at random on each page and count the words on those lines, we shall certainly get a satisfactory estimate of the average word length. That in fact was how fig. 1 was constructed.

Another method would be to stick a pin into the text at random places, and afterwards to take the words thus pierced. That would lead to a biased result, however, for the long words have a greater chance of being pierced than the short ones. Fig. 2, which was produced in this way, shows this very clearly. The average length is now 7.67 letters per word, quite a lot more than the value of fig. 1. Even so, we can still arrive along these lines at a correct result. The chance of piercing a word is clearly

proportional to the length of that word, and a statistical theory, making use of this datum, tells us that we must now compute the mean of the *reciprocal* word-length. The value found is 0.188 words per letter, and by taking the reciprocal of this mean we obtain a correct estimate for the actual average word-length:

$$\frac{1}{0.188} = 5.31 \text{ letters per word,}$$

which agrees very nicely with the value found from fig. 1.

This method finds practical application in traffic surveys for ascertaining the average distance travelled by motorists per journey. The system is to stop an arbitrary group of cars on the road and to ask the drivers where they come from and where they are going. The chance of a motorist being included in the sample is proportional to the length of his journey, and here again the above method of calculation must be applied.

Finally, there is a third method by which we can find the average word-length: we determine the average number of letters and the average number of words per full line of print, and then divide the first average by the second. The number of letters per line is found to be surprisingly constant, so that a few counts are sufficient to achieve reasonable accuracy. It is a simple matter to count the number of words in a line, and here an accurate result can be arrived at by making a large number of counts.

The three methods discussed are compared in *Table I*. Column B gives the average word-length, column C the 95% confidence interval, column D the width of this interval, and column E the total number of counts. The first two methods have roughly the same width of interval, but since the average word-length in the second method is greater,

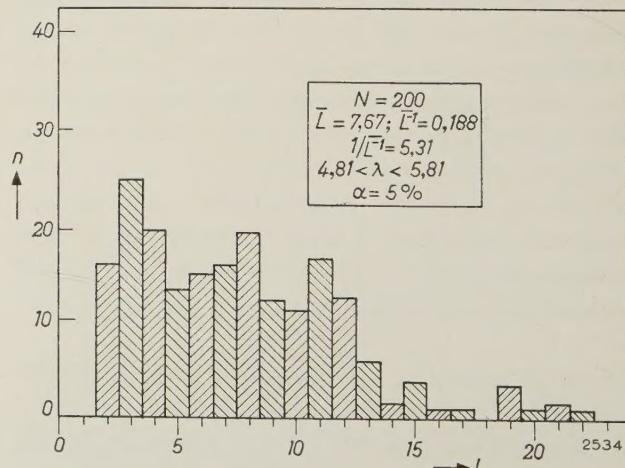


Fig. 2. Histogram of the lengths of 200 words, found by pricking the text blindly with a pin.

Table I. Comparison of the average word-lengths, \bar{L} , determined by three different methods. λ is the true average length.

Method	\bar{L} letters/word	Confidence interval for confidence level of 95%	Width of confidence interval	Total number of counts
I. Sampling by random numbers	5.34	$4.88 < \lambda < 5.80$	0.92	$200 \times 5.3 = 1060$
II. Sampling by pinpricks	5.31	$4.81 < \lambda < 5.81$	1.00	$200 \times 7.7 = 1540$
III. Average number of letters per line divided by average number of words per line	5.48	$5.22 < \lambda < 5.74$	0.52	$(7 \times 57.3) + (60 \times 10.5) = 1031$
A	B	C	D	E

many more counts were needed. Method II is thus decidedly less efficient than method I. In the third method the number of letters was counted in 7 lines, and the number of words in 60 lines, amounting altogether to 1031 counts, about the same as in method I. The resulting confidence interval, however, is only half as wide; method III therefore provides by far the most accurate result for the same number of counts. In this way, then, a statistical analysis enables us to decide on the method of investigation to be preferred.

The second question, concerning the accuracy to be aimed at, is countered by the statistician with the further question: For what purpose do you want to use the result? If it is simply a matter, as it is here, of collecting demonstration material for a lecture, no particular demands need be made on the accuracy, and the data given above are more than ample. But if we want to use the average word-length as a personal characteristic of the author, and to compare the discourses of various authors on that point, the position is entirely different.

From counts made (by the third method) on various published inaugural lectures I have not been able to discover any differences. That is not surprising, for the texts were all in Dutch, and differences, if any, would certainly be very slight. Much larger samples would be needed to find them.

The average word-length, then, is unsuitable as a characteristic of the literary efforts of different professors. In the average *sentence-length*, i.e. the average number of words per sentence, we may expect more variation. This characteristic may be found by a method similar to method III, that is by taking the product of the total number of printed lines and the average number of words per line, and dividing it by the total number of sentences. We can skip the details. An investigation of seven texts showed the overall average sentence-length to be

about 25 words, with individual averages varying from 22 to 32 words. Here the differences are quite marked.

One might try to pursue this line of inquiry even further. I fear, however, that the reader will lose patience at this point, feeling that a study of the length of words and sentences can only be of academic interest. To bring out more plainly the practical significance of the procedures described, I shall therefore apply a transformation to our problem. I now change professors into machines. The words uttered by professors shall be products brought forth by these machines, and the length of the words shall be a dimension characterizing the quality of these products, for example a diameter.

After this drastic transformation we still find ourselves faced with the same problems. Although the dimensions of the products are not as variable as the lengths of words, they are not all exactly the same; some variability is always present in mass production. Further, the number of products in many cases is so large that we cannot feasibly test them all, and we must again be content with a sample.

There are, however, marked differences between the machine problem and the professor problem. Whereas the average word-length was not subject to any special requirements, a nominal value is usually specified for the average diameter; deviations from that value tell us whether a machine is properly adjusted. Again, the words in a piece of prose are neatly arrayed side by side, and long and short words succeed one another with a certain regularity. But the products turned out by a machine are jumbled together in a heap, and so we can no longer apply method III to find the average diameter. Moreover, whilst no or only minor disparities were found in the average length of the words produced by different professors, and no variation is to be expected in the course of time,

the average diameters of the products from the several machines certainly will differ, and these values are also subject to variations over a short period as a result of wear. A single, elaborate investigation as represented in figs. 1 and 2 is out of place where machines are concerned; the appropriate method here is to take small samples at regular intervals, in order that prompt action can be taken if misalignment becomes too serious.

Nor do machines run on indefinitely. From time to time a breakdown occurs, or operations have to be stopped to make adjustments or feed in fresh material. We may thus imagine the lengths of the sentences to be transformed into time intervals during which the machines operate without interruption. Now, in any discourse short sentences generally improve the clarity and long sentences are better avoided, whereas long operating periods are just what we want from a machine; machines with a propensity for unduly short periods are consigned to the workshop for overhaul. Here again we have an evident difference.

Finally, we can go a step further and ask in how far discrepancies in the quality of the products, turned out by different machines, are attributable to the machines themselves, to the workers operating the machines, or to the batches of raw material processed. An answer to this and similar questions can best be supplied by an experiment, along the lines, for example, of a "Latin square", as represented in fig. 3. We choose, say, 4 machines, 4 operators and 4 batches of raw material. Each batch is divided into 4 portions and the experiment is arranged so that each batch is processed once on each of the 4 machines and once by each of the 4 operators, while each operator also works once on each machine. With an experimental design of this type,

	A	B	C	D
I	2	4	1	3
II	3	1	2	4
III	4	2	3	1
IV	1	3	4	2

2535

Fig. 3. Statistical design for an experiment, a "Latin square", used for investigating the influence of machines, labour and raw materials on the quality of a product.

A, B, C, D machines.

I, II, III, IV operators.

1, 2, 3, 4 batches of material.

the influence of the machine, the operators and the raw material can be separately determined by an appropriate analysis of the data — a typical instance of the way in which the statistician interweaves several factors in a single experiment. With the Latin-square design the information required is obtained with a minimum number of observations.

The reader will understand that experiments of this kind are not appropriate to the study of the lengths of words or sentences. One might ask, for example, in how far the average sentence-length, which — as we have seen — is characteristic of the individual professor, is also dependent on his subject matter. But an experiment in which a number of professors were each persuaded to write a number of inaugural addresses on different subjects, is out of the question.

The radical transformation I have applied thus demonstrates the point emphasized earlier: the statistical nature of the problems is not affected by such a transformation but the questions to which we seek an answer, and the conditions under which we must work, may be substantially altered. Each case, then, demands an individual approach to arrive at a satisfactory solution.

Finally, I should like to touch briefly on the question in what respect applied statistics differs from theoretical or mathematical statistics. There is reason for doing so in that the practitioners of these two branches of statistical science tend very much to keep their distance. A journal of such broad scope as the American "Annals of Mathematical Statistics", for example, prints many articles that are beyond the grasp of most applied statisticians, and give them the feeling that this is a special kind of mathematics for the sake of mathematics, but without any connection with the practical problems of daily life.

In our earlier count of 200 words we established a confidence interval for the average word-length λ with a confidence level of 95%. The question turns now on what degree of accuracy may be attached to this quantity. The applied statistician is interested only in the order of magnitude. For him a confidence level of 95% means that the risk of reaching a false conclusion is of the order of 5%; but it is of no serious consequence if the real risk is not 5% but 2% or 10%.

The mathematician, on the other hand, would require the risk to be exactly 5%, or, if that is mathematically unattainable, he requires 5% to be an upper limit to the risk; on no account should the risk be greater than 5%.

The science of statistics owes a great deal to this striving after mathematical rigour. It has led, for instance, to the development of the elegant and often very simple "distribution-free" methods, to which Van Dantzig and his pupils at Amsterdam have contributed so much³⁾.

But an exaggerated insistence on rigour has its practical disadvantages. It does not always result in simplifications; on the contrary, it produces a multiplicity of methods and formulae which may hamper rather than assist the effective application of statistical techniques.

What, then, it may be asked, is the real value of statistical analysis of observations? It does not, as I see it, lie in the exactness of the risks attached to conclusions. Statistical techniques provide us with internationally accepted standards for dealing with questions that arise in all research work. It is the international acceptance of uniform procedures that determines the real value of statistical methods. It leads to the economical formulation of the results and interpretation of scientific investigations. Undoubtedly international acceptance is more likely when the risks in a statistical procedure are clearly established. But simplicity and uniformity of methods and formulae are equally essential. If the pursuit of scrupulous accuracy produces a multitude of methods and complicated formulae, uniformity goes begging and it may therefore be preferable to relax the demands on exactness.

The difference in viewpoint between the theoretician and the practitioner has far-reaching effects. If we want to analyse the separate influences of operator, machines and raw materials on the quality of the products, using a Latin square as in fig. 3, the assumption underlying the experiment is that these three factors operate *independently*. We assume, for example, that the differences between the operators are always the same for any given combination of materials and machines. But if each operator is specially accustomed to his own machine, we are not entitled to interchange men and machines freely, and the Latin square is no longer the proper design for the experiment.

Every statistical experiment is thus based on a *model*, that is to say a mathematical formula which indicates in what way the factors that are varied influence the result. The strict theoretician wants this model to be completely established before the

experiment is embarked upon, and he will not allow any modification of the model on the grounds of the experimental results; for any such subsequent alteration would mean that the risks inherent in the statistical conclusion could no longer be exactly evaluated and would certainly be increased. This the mathematician will not allow.

The implication, however, would be that statistical methods were only applicable in experiments whose outcome was a foregone conclusion and whose purpose was merely to establish numerically the parameters occurring in the model. In my view the experimenter can never accept such a standpoint. The object of many experiments is exploratory: it is often those experiments that yield unexpected results that are the most valuable. If the theoretical statistician washes his hands of such experiments as being beyond the scope of a statistical analysis, the experimenter for his part will have little interest in statistics.

Another way of trying to circumvent the choice of confidence limits is to interpret the problem "econometrically". It is then assumed that a decision has always to be taken at the end of an investigation; for example, a sampling inspection to determine the quality of a batch of products ends in a decision to accept the batch or to reject it. Considered statistically, there is always some risk of taking a wrong decision, that is of rejecting a good batch or of accepting a bad one. We can reduce these risks, however, by taking a larger sample. Now if we know the economic damage caused by wrong decisions, we can weigh the risk of damage against the costs of the inspection, and settle the size of the sample so as to minimize the average total costs.

But this approach leads to endless complications. The costs which a wrong decision may entail are difficult to determine and are subject to a high degree of uncertainty. They depend, for instance, very markedly on available stocks if the objects concerned are components intended for further production. Where stocks are plentiful, we need have no scruples about rejecting bad or doubtful lots, but if stocks are nearly exhausted, the rejection of a lot may soon force a manufacturer to stop production for lack of material. In a short time, then, the damage due to a wrong decision may vary over a very wide range.

When a supplier subjects a batch of finished products to sampling inspection, he does so partly to ensure that his good name will not suffer from the delivery of an inferior consignment. But the economic value of a good name and the harm done to it can be little more than a conjecture.

³⁾ See e.g. the bibliography in Statistica neerlandica 13, 432, 1959 (No. 4).

"Distribution-free" methods (also termed, less appropriately, "non-parametric" methods) are those which are not based on the assumption that the error distribution is necessarily normal or gaussian.

Furthermore, we must know what happens to rejected batches of products. They may, for example, be returned to the supplier, or subjected to a more thorough inspection, or sold at a reduced price, or written off as a total loss. It also makes a difference whether we consider the costs from the standpoint of the customer or of the producer, or of both together.

Here again, the conflict between theory and practice is clearly apparent. If 5 out of a sample of 100 products, i.e. 5%, are found defective, we can assert with a confidence of 95% that the lot from which the sample was taken will contain at least 1.9% and at the most 10.7% defective products. This is a purely statistical assertion, which is arrived at with the aid of an elementary table and leaves us free to adopt whatever course of conduct we may think reasonable. However, as soon as we attempt to define that reasonable course of conduct more exactly on a mathematical basis, we are up against not one but a multiplicity of situations, and we soon find ourselves in a maze of theories through which only an experienced theoretician can find his way. These theories may be successfully applied in specific cases, but they lack the general usefulness and wide compass of the confidence interval.

Between these rocks the applied statistician must learn to navigate. His starting point should be that the function of statistics in experimental science is that of a servant. The experimental scientist is guided by the stepping-stones of experience and intuition. That, in my opinion, is essentially characteristic of all research, and we have to accept the fact that experience and intuition are vague concepts that cannot be pinned down in mathematical formulae and numerical constants.

If the precise formulation of statistical theories is pursued too far, one is driven to express the vague concepts of experience and intuition in terms of exact numerical parameters. Doubtless the theories thus produced may be interesting, but I fear they will generally prove sterile in their application.

The applied statistician, then, is bound to accept the experience and intuition of the experimenter as his working basis, though at the same time preserving a critical attitude. For experience and prejudice are often intermingled and difficult to disentangle; and where prejudice is suspected, verification is required.

When a statistician is consulted on the solution of a practical problem, his first job is to find out, from a discussion with his principal, what may be regarded as well-founded experience and where

preconceived opinions may have crept in. When it comes to interpreting the observational data, constant cooperation is called for between statistician and experimenter. If the latter is led by the observations to modify his views on the nature of the phenomena investigated, the statistician must adapt his model accordingly. But he should be aware of the dangers of such a procedure, and warn against it where necessary. Practised in this way, applied statistics constitutes a fascinating and valuable chapter in the development of experimental science.

BIBLIOGRAPHY

Some publications by the author on the subject of applied statistics:

Beispiele zur Anwendung statistischer Untersuchungsmethoden in der Industrie, *Mitt. Bl. math. Stat.* 5, 211-229, 1953.

De betekenis van de statistiek voor de ontwikkeling van de experimentele wetenschap, *Sigma* 1, 55-58, 1955.

Naar efficiënte experimenten, *Statistica neerl.* 9, 7-25, 1955. Experimental design in industry, *Biometrics* 11, 257-286, 1955.

Infusing statistics into industry, *Bull. Inst. Intern. Stat.* 35, 433-443, 1957.

Statistiek en experiment, *Statistica neerl.* 12, 119-130, 1958.

De recente ontwikkeling in proefopzetten met kwantitatieve factoren, *Statistica neerl.* 12, 201-212, 1958.

Summary. In many fields of inquiry, especially those concerned with living beings, "exact" observations are not possible and it is necessary to investigate the effect of several factors at the same time. This has led to the design of experiments on a statistical basis, in which several factors may be varied simultaneously and which require an appropriate statistical analysis for their interpretation (R. A. Fisher). In this his inaugural lecture as professor at the Eindhoven Technische Hogeschool, the author advances the view that the introduction of statistical principles is bringing about a fundamental change in experimental science in general. Though the statistical principles are always the same: the methods adopted depend on the nature of the investigation. In agriculture, medicine, sociology, industry, etc., quite different techniques may be needed. To illustrate the statistician's approach, the author inquires into the average length of the words and sentences in his own discourse and in similar discourses by others. Three methods for determining the average word-length are discussed and compared. Next, the same principles are applied to problems of quality control of machine-made products. Though the problems are similar, the circumstances are different, particularly in regard to conditions (e.g. variation in time), aims and experimentation possibilities. Finally, the author considers the marked differences between applied and theoretical statistics. The true value of statistical methods is that they provide universally accepted standards for the conduction and interpretation of investigations. It is of great value that with statistical procedures the inevitable risk of drawing a wrong conclusion is known. For the applied statistician, however, only the order of magnitude of this risk is of interest, while the mathematician requires that the risk shall be known exactly. Experimental science proceeds on the basis of experience and intuition, vague concepts that cannot be brought under the rules of precise mathematical formulae and constants. The applied statistician should take this fact as his point of departure and should not attempt to achieve a precision which does not correspond to the actual conditions of experimentation. Applied in this way statistics will prove to be a valuable aid to the experimenter.

WAVEGUIDE EQUIPMENT FOR 2 mm MICROWAVES

I. COMPONENTS

by C. W. van ES, M. GEVERS and F. C. de RONDE.

621.372.8

Microwaves of millimetre wavelengths are an important tool in plasma research and in research on the solid state. Special equipment is therefore needed for the measurement of frequency, phase, power, absorption and reflection coefficient in the microwave region. The article below is part I of a survey of equipment developed at Philips for 2 mm waves. The equipment consists partly of scaled-down versions of equipment for longer waves, and partly of entirely new designs. Part II, to be published later, will describe various set-ups for microwave measurements at wavelengths of 2 mm.

Microwave techniques are nowadays the subject of considerable interest, for general scientific reasons as well as in connection with the development of technical applications. The main branches of science in which microwave techniques are of importance are plasma physics, microwave gas spectroscopy and solid-state research.

In plasma physics, interest is focused on the electron density and temperature of plasmas. The electron density may be derived from the propagation characteristics of millimetre waves which are passed through the plasma. One way of determining the temperature is to compare the energy radiated by the plasma in a certain millimetre wave region with that of a radiation source of known intensity (noise generator) in the same waveband. Such measurements are of particular importance in experiments relating to nuclear fusion¹⁾.

Microwave gas spectroscopy is concerned in particular with the study of molecular rotational energy. In principle a microwave gas spectrometer consists of a microwave generator, a vacuum-tight waveguide filled with the gas under investigation, and a detector. The equipment measures the absorption occurring at specific frequencies²⁾.

A third application is in the field of solid-state research. When an electromagnetic wave is directed upon a solid, such as germanium or silicon, in which a magnetostatic field is present perpendicular to the electric field, cyclotron resonance will occur at certain frequencies³⁾, yielding information, e.g., on the charge carriers. Millimetre waves are also used in research on superconductivity⁴⁾.

For investigations of this kind, suitable equipment is needed for generating the microwaves, for propa-

gating them and performing measurements with them. As regards microwave generators, the Philips 4 and 2.5 mm reflex klystrons have proved very satisfactory⁵⁾. The reflex klystron used for the present range of equipment is the 4 mm DX 151, the wavelength being reduced to 2 mm by means of a frequency doubler (point-contact diode).

For propagating the 2 mm waves, waveguides of rectangular cross-section are generally used (0.83×1.66 mm or $0.0325'' \times 0.065''$).

The measurements involved usually relate to frequency, phase, power, absorption and reflection coefficient. Components for measuring these quantities in the centimetre wave region were described in this journal ten years ago⁶⁾. Since then, comprehensive series of components have been developed for wavelengths of 3 cm, 8 mm and 4 mm. At the latter wavelength the utmost care is needed to keep the dimensions within the required tolerances. Even so, it has proved possible, with the same relative tolerances, to develop a range of components for 2 mm waves⁷⁾, the more important of which will be reviewed in this article. Some are scaled-down versions of comparable components for longer waves; others differ mechanically from the familiar designs, and three are entirely new, namely: the *PIN modulator*, the *variable impedance* and the *rotary directional coupler*.

All components for 4 and 2 mm waves are fitted with claw flanges (to be discussed presently), by

¹⁾ M. A. Heald, Microwave measurements in controlled fusion research, Nat. Conv. Rec. Inst. Radio Engrs. **6**, No. 9, 14-18, 1958.

²⁾ W. C. King and W. Gordy, One-to-two millimeter wave spectroscopy, Phys. Rev. **93**, 407-412, 1954.

M. Cowan and W. Gordy, Further extension of microwave spectroscopy in the submillimeter wave region, Phys. Rev. **104**, 551-552, 1956.

³⁾ C. J. Rauch, J. J. Stickler, H. J. Zeiger and G. S. Heller, Millimeter cyclotron resonance in silicon, Phys. Rev. Letters **4**, 64-66, 1960 (No. 2).

⁴⁾ M. A. Biondi and M. P. Garfunkel, Millimeter wave absorption in superconducting aluminum, Phys. Rev. **116**, 853-867, 1959 (No. 4).

⁵⁾ B. B. van Iperen, Reflex klystrons for wavelengths of 4 and 2.5 mm, Philips tech. Rev. **21**, 221-228, 1959/60 (No. 8). Another method of generating millimetre waves is described by M. Yéou-Ta, Carcinotrons du type O fonctionnant sur une longueur d'onde de 2 mm, Onde électr. **39**, 789-794, 1959 (No. 391).

⁶⁾ A. E. Pannenborg, A measuring arrangement for waveguides, Philips tech. Rev. **12**, 15-24, 1950/51.
A. E. Pannenborg, Some aspects of waveguide technique, Communic. News **11**, 65-75, 1950.

⁷⁾ We are particularly indebted to the instrument-maker J. van Ostade for his contribution to this work.

which their respective waveguides are joined together. This construction makes for compactness and permits ready assembly and disassembly of microwave systems.

Generation, detection and modulation of 2 mm waves

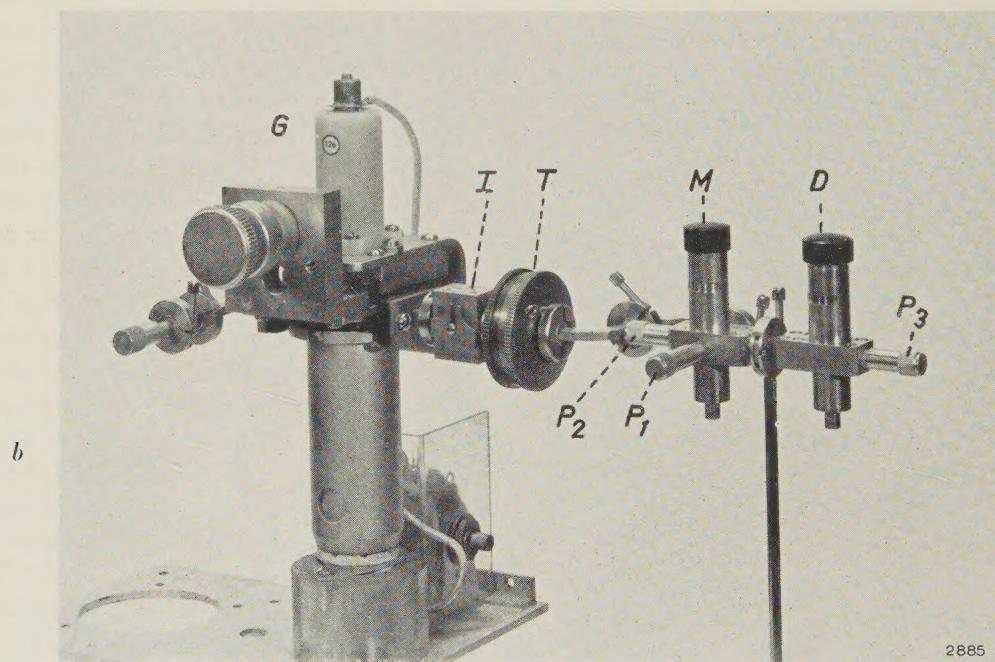
The generator, the frequency multiplier and the detector

As mentioned above, the 2 mm waves are generated in our equipment by doubling the frequency of a reflex klystron, type DX 151, designed for 4 mm waves. Mounted between the klystron and the frequency multiplier are an isolator *I* and a tuner *T* (fig. 1). The isolator prevents feedback from the load to the generator, whose frequency and power output would otherwise vary with the load. The principle of this isolator depends on gyromagnetic resonance⁸⁾. The tuner, combined with a shorting plunger *P*₁, ensures that virtually the whole power output of the klystron reaches the frequency multiplier. The design of the tuners used will be described later.

⁸⁾ H. G. Beljers, Resonance isolators for millimetre waves, Philips tech. Rev. 22, 11-15, 1960/61 (No. 1).

The frequency multiplier consists of a point-contact diode situated at the junction between the 4 mm waveguide *1* and the 2 mm waveguide *2* (fig. 2). Projecting through these two waveguides is a tungsten catswhisker *3* which, together with a silicon crystal *4*, forms the point-contact diode. The lower end of the catswhisker is welded to an insulated pin *5*. The other end is etched to a fine point. The silicon wafer, which can be finely adjusted in the vertical direction by a differential screw, is brought into contact with the whisker. The part of the catswhisker running through the two waveguides acts as a coupling probe. The resistance of the diode (and also the capacitance) depends on the applied voltage⁹⁾. As a result of this non-linearity the current induced in the coupling probe by the incident 4 mm waves contains a component with twice the frequency. This produces a 2 mm wave in

⁹⁾ H. C. Torrey and C. A. Whitmer, Crystal rectifiers, M. I. T. Radiation Lab. Series, 15, McGraw-Hill, New York 1948, in particular pages 97-107, Rectification at high frequencies. R. S. Ohl, P. P. Budenstein and C. A. Burrus, Improved diode for the harmonic generation of millimeter and submillimeter waves, Rev. sci. Instr. 30, 765-774, 1959 (No. 9).



2885

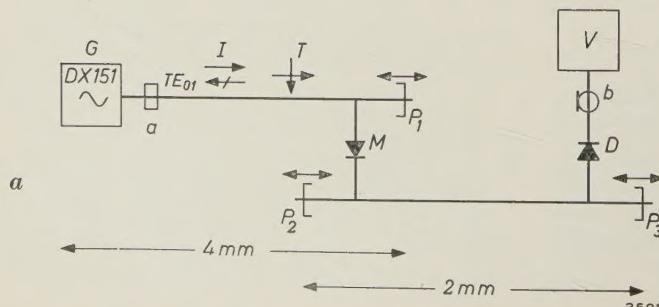


Fig. 1. Block diagram (a) and photograph (b) of a set-up for generating 2 mm waves. *G* generator of 4 mm waves (reflex klystron DX 151). *I* isolator. *T* tuner. *P*₁, *P*₂, *P*₃ adjustable waveguide plungers. *M* frequency multiplier. *D* detector for 2 mm waves, with millivoltmeter *V*. The sign *a* means "rectangular waveguide" and the sign *b* means "coaxial line".

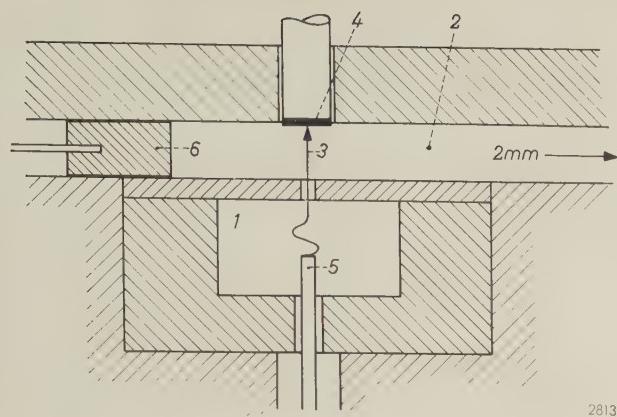


Fig. 2. Cross-section of frequency multiplier. 1 is the 4 mm waveguide, 2 the 2 mm waveguide. 3 cat's whisker and 4 silicon crystal, together forming a point-contact diode. 5 insulated pin carrying the cat's whisker. 6 shorting plunger in the 2 mm waveguide (P_2 in fig. 1).

the upper waveguide. Fig. 3 is a photograph of the frequency multiplier.

To verify that the equipment is functioning properly up to the frequency multiplier, the latter can be used temporarily as a 4 mm-wave detector. For this purpose a DC voltmeter (e.g. type GM 6020 electronic millivoltmeter) is connected between the central pin 5 and earth — i.e. in parallel with the point-contact diode. If the meter reading is satisfactory, a crystal detector D (fig. 1) is then connected to the 2 mm waveguide for the purpose of detecting the 2 mm waves. The detector is virtually identical with the frequency multiplier, except of course that it has no 4 mm waveguide.

The pivoting screw tuner

The matching for the frequency multiplier and the 2 mm detector is effected by waveguide plungers (P_1 , P_2 and P_3 in fig. 1) in conjunction with a tuner of special design. At longer wavelengths a tuner of the sliding screw type is ordinarily used. This consists of a metal stub 1 (fig. 4a), whose depth of penetration into the waveguide 2 can be adjusted by turning the screw, and which can also be moved along the waveguide, in a slot, over a distance of about $\frac{3}{4}\lambda_g$ (the wavelength in the waveguide, λ_g , is greater than the corresponding wavelength λ in free space).

The relation between λ_g and λ is given by:

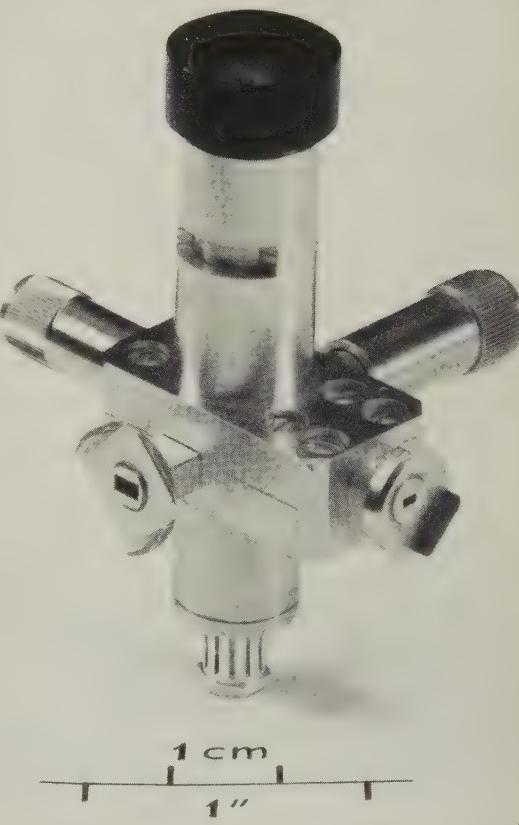
$$\frac{1}{\lambda^2} = \frac{1}{\lambda_g^2} + \frac{1}{\lambda_c^2},$$

where λ_c is the cut-off wavelength. (See p. 17, formula (2) of the first article mentioned in reference 6.) For the mode used in the waveguide — the TE_{01} mode — λ_c is equal to twice the width a of the rectangular waveguide. For $\lambda = 2.00$ mm and $\lambda_c = 2a = 3.32$ mm, the wavelength λ_g is 2.51 mm.

The depth to which the stub penetrates into the waveguide regulates mainly the extent to which the stub reflects the wave; displacement of the stub along the guide varies the phase of the reflected wave. Reflections in the equipment can thus be cancelled by generating with the sliding screw tuner a reflection of equal magnitude but opposite phase. In this way a travelling wave is produced.

The 2 mm-wave screw tuner is based on a new design. The stub does not slide along the waveguide, but turns about an axis 3 (fig. 4b), so that it describes a pivoting motion, hence the term "pivoting screw tuner". In the millimetre wave region this design has special mechanical and electrical advantages:

- 1) The objection to the sliding screw tuner, that the stub may move slightly up and down as a result of surface irregularities, is entirely avoided here; there is less friction and adjustment is smoother.
- 2) The axial length of the tuner, from flange to flange, is limited to 12 mm, as a result of which



2804

Fig. 3. The frequency multiplier. Foreground, left: connection for the 4 mm waveguide; right, connection for the 2 mm waveguide. The arms on the opposite side are fitted with waveguide plungers. The vertical arm contains the point-contact diode, and is surmounted by the adjusting knob. At the bottom is the coaxial connection for the millivoltmeter. The conversion loss is 15 to 20 dB.

the pivoting screw tuner causes particularly little attenuation (no more than 0.1 dB).

At first sight it might seem a drawback that, since the stub is pivoted about the axis 3, its depth as well as its position varies in the waveguide, implying

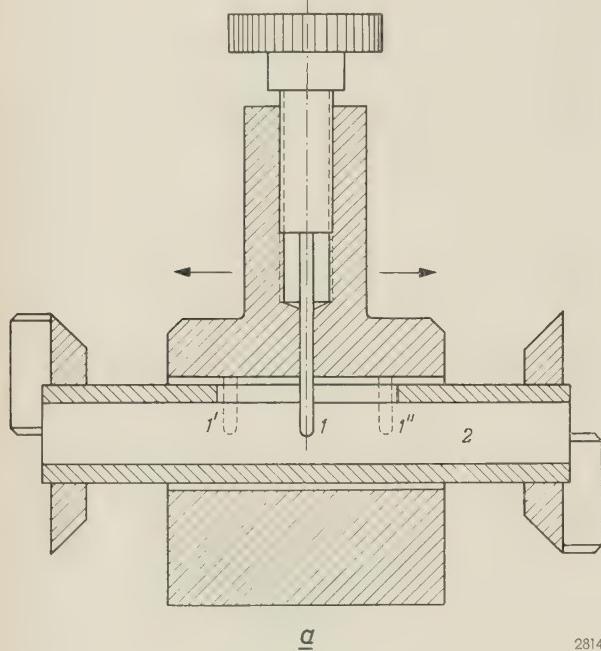
that the amplitude of the reflected wave will also vary slightly with the phase. This effect is so small, however, that it gives no trouble in practice.

A photograph of the pivoting screw tuner appears in fig. 5.

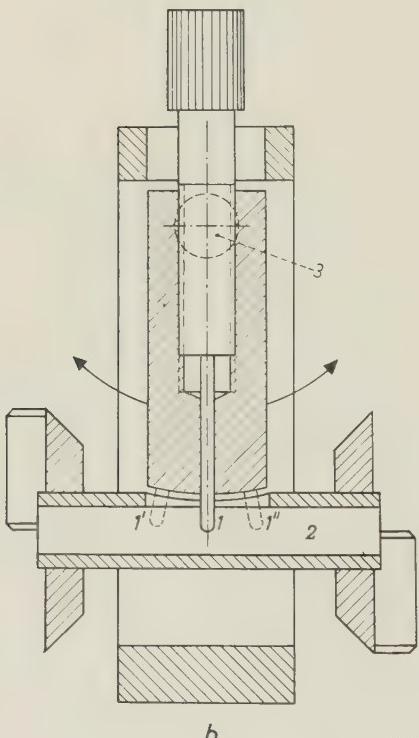
Methods of modulation

At an input power of, say, 100 mW the frequency multiplier delivers a 2 mm output of up to 2 mW. To detect the 2 mm wave at various points in the equipment, where the power may be much lower still, a highly sensitive method of measurement is therefore required.

For this purpose, in accordance with a widely used principle, the wave is modulated in amplitude and a selective indicator is employed behind the detector. The indicator consists of an amplifier tuned to the modulation frequency (e.g. 800 c/s), a low-frequency detector and a DC voltmeter. With a modulation depth of 100% the meter will then give a full scale deflection at a microwave power of the order of 1 μ W. The sensitivity can be raised still further by the use of synchronous detection; in that case, to keep the noise level down, the modulation frequency should be fairly high (e.g. 8000 c/s).

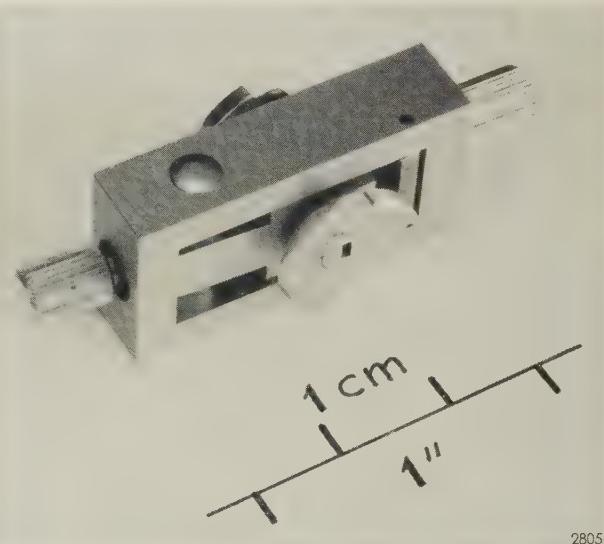


2814



2815

Fig. 4. Schematic representation of screw tuners.
 a) Sliding screw tuner. The dept of the stub 1 in the waveguide 2 can be adjusted by turning a screw, and its position along the waveguide can be adjusted between I' and I''.
 b) Pivoting screw tuner. Here too the stub 1 can be screwed in and out, but instead of sliding lengthwise it now pivots about spindle 3. The extreme positions are again I' and I''.



2805

Fig. 5. Pivoting screw tuner for 2 mm waves. Left, the knob for screwing the stub in and out of the waveguide; right, the knob for altering its position (by pivoting about 3 in fig. 4b). This component is capable of effecting a match in all cases encountered in practice. The attenuation is no more than 0.1 dB.

We shall now consider briefly some of the modulation methods used. The most common method is to periodically vary the repeller voltage V_r of the reflex klystron between the value for optimum oscillation and a value at which the klystron does not oscillate. An objection to this method is that the oscillation

frequency also depends on V_r ¹⁰⁾; to avoid unwanted frequency modulation one would have to give V_r a square waveform. In practice, however, it is very difficult to make the edges sufficiently steep and to keep V_r constant enough in the oscillation interval. Some frequency modulation therefore remains, which can be particularly undesirable where bridge circuits are involved.

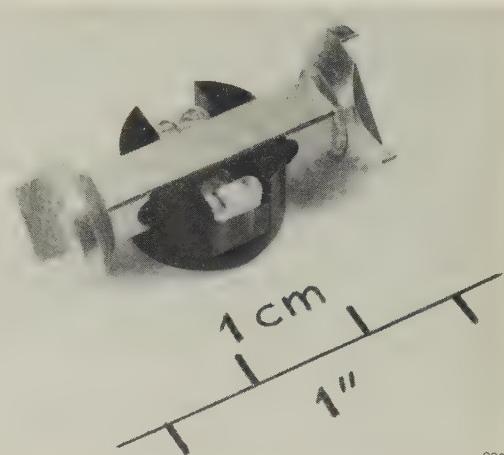
In our case, where a frequency multiplier is used, the modulation can be simply effected, without frequency modulation arising, by applying a low-frequency alternating voltage across the diode of the frequency multiplier, in series with any biasing voltage present. An alternating voltage of about 1 V is quite sufficient to modulate the 2 mm wave 100% in amplitude.

Without a frequency multiplier, pure amplitude modulation is possible by making use of a new device called the *PIN* modulator¹¹⁾. This is an electrically controlled attenuator which is inserted in the waveguide. It consists of a germanium wafer in which a *P* region and an *N* region are separated by an *I* region of pure germanium, possessing extremely low "intrinsic" conductivity. Only the *I* region is contained inside the waveguide. In the normal state it behaves like a low-loss dielectric. When, however, a voltage is applied between the *P* and *N* regions in the forward direction, holes and electrons are injected into the *I* region, which thereby becomes an absorbing medium. In this way, using a control current of 15 mA (control power approx. 10 mW), the losses in the *I* region can be made so high as to increase the attenuation by 25 dB. If an alternating control current is used, the wave is thus modulated in amplitude. A *PIN* modulator for 2 mm is illustrated in fig. 6.

Waveguide and claw flanges

Except in the rotary attenuator, the variable impedance and the rotary directional coupler, where circular waveguides are used for reasons presently to be discussed, the waveguides in the various components are all rectangular.

The rectangular waveguides are constructed of a U-section, made by milling a groove in rectangular bar stock, to which a flat cover plate is screwed or soldered. In most cases the material used is brass;



2806

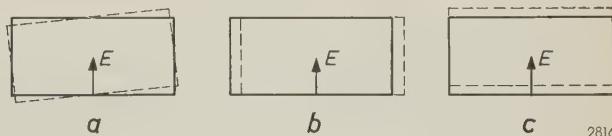
Fig. 6. *PIN* modulator for 2 mm waves. The attenuation is adjustable from about 5 dB to about 30 dB. The modulation frequency at which the modulation depth has dropped by 3 dB is approximately 50 kc/s.

the waveguides are gold-plated inside and out in order to keep the surface properties constant. Silver is used only where minimum attenuation is required.

The internal dimensions of the 2 mm waveguides are 0.83×1.66 mm, with a tolerance of 0.01 mm. The outside dimensions are determined solely by the mechanical considerations applicable to the part in question. The internal cross-section must be uniform, otherwise a discontinuity can arise between coupled components, giving rise to reflections. But even when the cross-sections are uniform, reflections may still occur if the walls of the coupled sections of waveguide are not well-aligned (fig. 7a, b and c)¹²⁾. Such misalignment points to shortcomings of the waveguide coupling (the flanges and the means by which they are fixed together). An essential requirement of a waveguide coupling is that the walls should be in true alignment with one another.

Other desirable properties of a waveguide coupling are:

- 1) The axial length should be small in order to mini-



2816

Fig. 7. Excessive tolerances between the flanges mutually or of the flange with respect to the waveguide may result in misalignment of the walls, e.g. the walls may be slightly turned in relation to one another (case a), or displaced at right angles to the electric field E (case b), or displaced in the direction of E (case c).

¹⁰⁾ See article by Van Iperen in reference ⁵⁾, page 224 (fig. 6 and note ⁴⁾).

¹¹⁾ F. C. de Ronde, H. J. G. Meyer and O. W. Memelink, The *PIN* modulator, an electrically controlled attenuator for mm and sub-mm waves, Trans. Inst. Radio Engrs. on microwave theory and techniques MTT **8**, 325-327, 1960 (No. 3). See also Dutch patent application No. 229 531 of 11th July 1958.

¹²⁾ U. von Kienlin and A. Kürzl, Reflexionen an Hohlleiter-Flanschverbindungen, Nachr.techn. Z. **11**, 561-564, 1958.

mize attenuation and frequency-dependence¹³⁾. This applies both to the flange and to the space between flange and component. The construction should be such that no tools need be used between flange and component; for this reason the joint should preferably be secured from a direction perpendicular to the waveguide axis.

- 2) The losses should be low and also reproducible, i.e. the losses should not be dependent on the force applied to draw the two flanges together. This force should preferably be applied by means of no more than one screw. The construction illustrated in fig. 8, where the flanges are fixed together by four screws, is not only unpractical but fundamentally wrong, because it is not possible to check whether the screws have all been tightened equally.
- 3) The coupling should be "universal", i.e. it should be possible to reverse the components in the

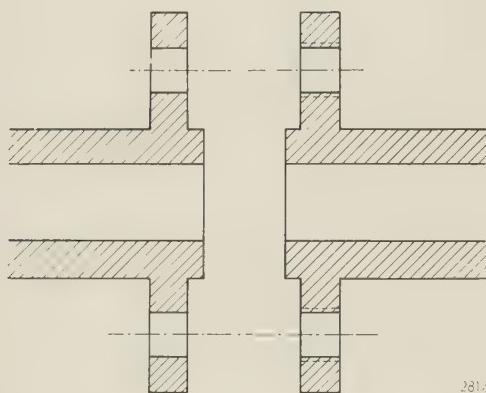


Fig. 8. Conventional waveguide joint. The flanges are joined by four bolts. The objections to this system are:

- 1) It is impossible to check whether the bolts have all been equally tightened.
- 2) The waveguide section between flange and component is necessarily relatively long, owing to the presence of the clamping bolts.

¹³⁾ The frequency range in which a rectangular waveguide operating in the TE_{01} mode can be used is from about 1.2 to 1.9 times the cut-off frequency $f_c = c/2a$ (c = velocity of light, a = width of waveguide).

waveguide system and also to connect them in reversed positions about the axis of the waveguide.

A flange that meets all these requirements is the *claw flange*, which is fitted to all our 2 mm components. Fig. 9 shows two claw flanges and the clamping ring consisting of two hinged halves, by which the flanges are drawn tightly together by means of a single screw perpendicular to the waveguide axis. The whole coupling — the two flanges and the clamping ring — has a total length of twice the flange thickness, namely 5 mm. As appears from the various photographs in this article, the flanges are fitted close up against the components, making for a particularly compact assembly.

Description of some 2 mm components

Some components of the microwave equipment will now be discussed in somewhat more detail, viz. three adjustable components, provided with dials: the rotary attenuator, the variable impedance and the rotary directional coupler.

The rotary attenuator

It is sometimes necessary in microwave measurements to adjust the power level of the transmitted wave. There is also often a need to introduce a known attenuation. In both cases, use is made of a variable attenuator. One of the oldest types is the vane attenuator, the principle of which is illustrated in fig. 10. An insulating plate V (the vane), coated with an absorbent resistive layer, is inserted in the waveguide, to an adjustable depth, parallel to the direction of propagation and to the electric field E . The absorption caused by the vane is greater the deeper the vane extends into the waveguide and the greater its axial length l with respect to the wavelength. The drawback of the latter fact is that the attenuation is partly dependent on frequency. If l is not large with respect to the wavelength, the attenuation will moreover be dependent on the standing-wave ratio in the waveguide.

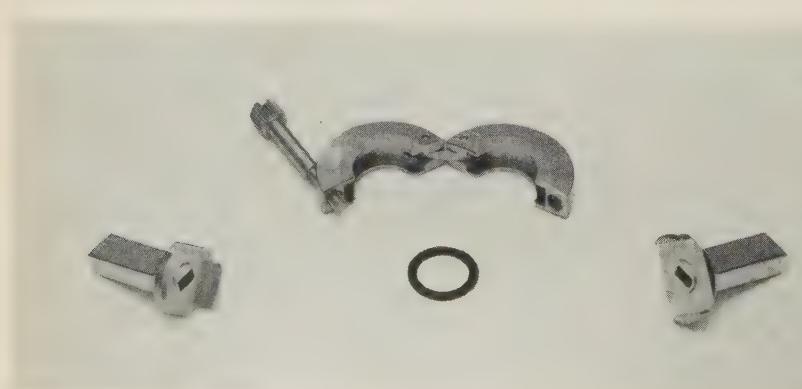


Fig. 9. Left and right, sections of waveguide with claw flanges; top centre, the hinged clamping ring which joins the flanges together; below, a rubber ring. (True size.) The clamping ring is tightened round the flanges with a single screw, perpendicular to the axis of the waveguide. The rubber ring ensures an airtight seal; the air in the whole waveguide system can then be put under pressure to prevent the entry of moisture.

The modulus of the reflection coefficient of the claw-flange coupling is smaller than 0.005.

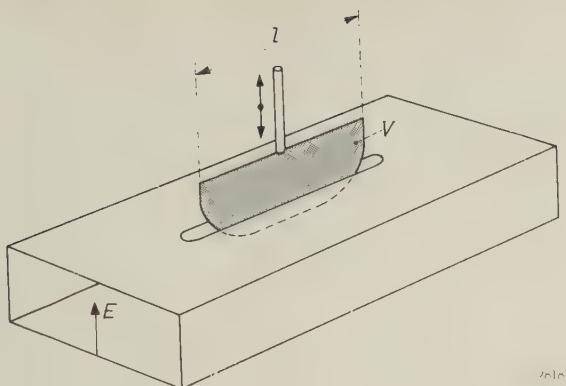


Fig. 10. Principle of the vane attenuator. An absorbent vane V , parallel to the electric field E , can be inserted to a variable depth in the waveguide through a narrow slot. The attenuation increases with the depth of insertion, but also depends on the ratio of the length l of the vane to the wavelength, i.e. it is frequency-dependent.

These drawbacks are entirely absent in the *rotary attenuator*, a device invented in America during the war by Bowen¹⁴⁾. A further advantage of this attenuator is that it is an absolute instrument, that is to say the scale giving the attenuation in dB obeys a mathematical law, and thus can be engraved at once to the instrument without previous calibration.

A schematic cross-section of the rotary attenuator is given in fig. 11a. The ends 1 and 2 are rectangular,

and the middle section 5-7-6 is circular. The transition between the rectangular and the circular cross-sections takes place in the regions 3 and 4. In the circular section the centre piece 7 is rotatable about the waveguide axis. In each of the (fixed) regions 5 and 6, an absorbing vane (8, 9) is mounted perpendicular to the electric field, i.e. parallel to the long sides of the rectangular waveguide section (fig. 11b). The central, rotatable part 7 of the waveguide also contains an absorbing vane, 10. This can be rotated from a position parallel to 8 and 9 to a position perpendicular thereto; as we shall see, the extreme positions correspond to the minimum and maximum attenuation, respectively.

The attenuation as a function of the angle α between the plane of the vane 10 and that of the vanes 8 and 9 is found as follows. The electric field E of a wave entering at 1 is perpendicular in the fixed section 5 to the vane 8; the wave here therefore passes through unattenuated. In the rotary section 7 we resolve E into a component $E \sin \alpha$ in the plane of vane 10 and component $E \cos \alpha$ perpendicular thereto (fig. 11c). The component $E \sin \alpha$ is entirely absorbed in the vane, leaving only the component $E \cos \alpha$. Finally, in the fixed section 6 the field is $E \cos \alpha$, whose component $E \cos \alpha \sin \alpha$, in the plane of 9, is absorbed by this vane so that only component $E \cos^2 \alpha$, perpendicular to 9, remains (fig. 11d).

The emergent wave thus has an electrical field-

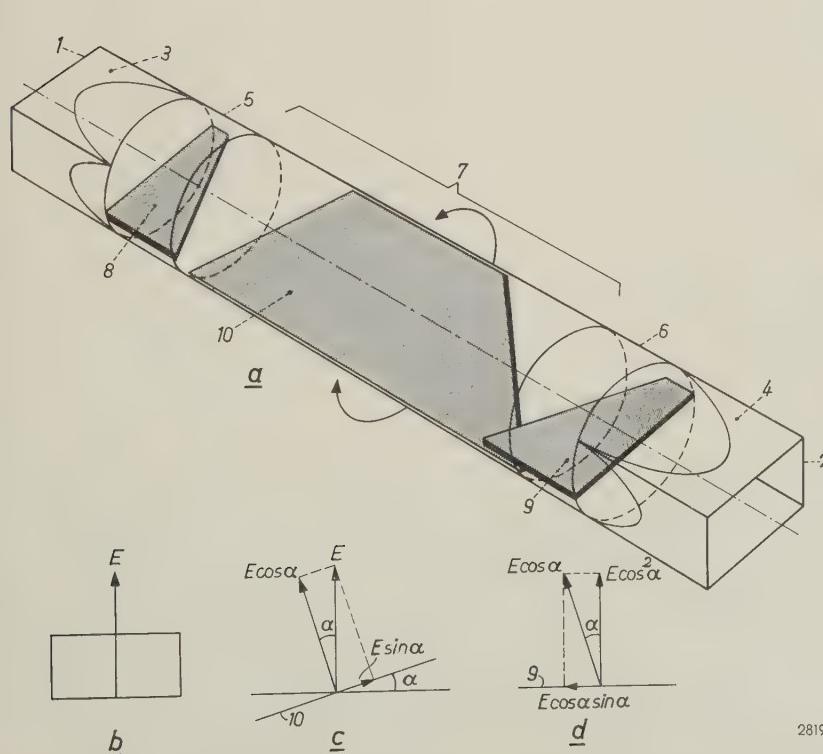


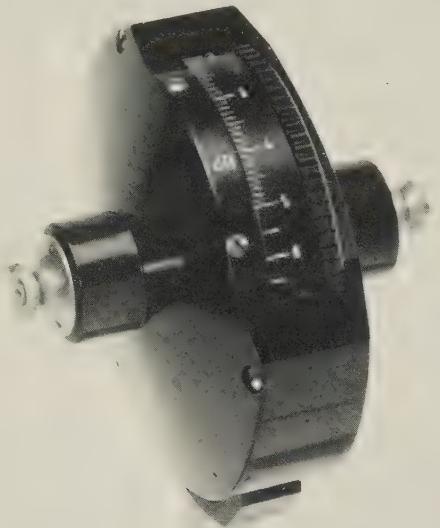
Fig. 11. a) Schematic diagram of the rotary attenuator. 1, 2 rectangular ends for connection to rectangular waveguides. 3, 4 transitions to circular waveguide. 5, 6 fixed sections of circular waveguide, 7 rotary section. 8, 9 and 10 vanes. b) Rectangular cross-section at 1 and 2 (long sides parallel to the vanes 8 and 9). E electric field.

c) The component $E \sin \alpha$ of the electric field of a wave entering at 1 is absorbed in vane 10, leaving only the component $E \cos \alpha$ (α is the angle made by the plane of 10 with that of the fixed vanes 8 and 9). d) In section 6 the component $E \cos \alpha \sin \alpha$ of the remaining wave is absorbed by vane 9, leaving only the component $E \cos^2 \alpha$. The total attenuation is thus $\cos^2 \alpha$, or $-20 \log \cos^2 \alpha$ dB.

strength which is $\cos^2 \alpha$ times that of the incident wave. The attenuation is therefore solely a function of the angle α and amounts to $-20 \log \cos^2 \alpha$ dB. Without calibration, then, a dial can be fitted whose scale gives an absolute reading of the attenuation.

It is important that the vanes 8, 9 and 10 should reflect as little energy as possible. For this purpose they are tapered at the ends to form trapezia and are extremely thin, consisting of metallized slivers of mica. The thickness of the metal coating is small compared to the skin depth, so that the vane acts as a resistive layer which strongly absorbs the wave. Due to the tapering of the ends and the thinness of the vanes, the reflection coefficient of the attenuator is less than 0.02.

The 2 mm rotary attenuator is illustrated in fig. 12.



2808

Fig. 12. Rotary attenuator for 2 mm waves. The attenuation, which is independent of the frequency, can be read on a scale covering a range from 0 dB to 50 dB. The insertion loss is 3.5 dB. The standing-wave ratio is less than 1.05.

The variable impedance

At the end of a waveguide there is generally some reflection of energy⁶). The incident and the reflected waves give rise to a standing wave, as a result of which the electric field-strength in the waveguide shows a minimum (E_{\min}) at certain places and a maximum (E_{\max}) at others. The standing-wave ratio, E_{\max}/E_{\min} , is a measure of the mismatch (the extent to which the terminating impedance differs from the characteristic impedance), and in many

cases this is an important indication of the quality of a microwave component.

In the centimetre range the standing-wave ratio is measured in waveguides in the same way as in lecher lines, that is with a standing-wave indicator. In the millimetre wave range, however, it is difficult to produce a standing-wave detector having the required accuracy. Instead of the standing-wave ratio we can measure the *reflection coefficient*, which is simply related to the standing-wave ratio (see page 18 of the first article mentioned in reference⁶). With the aid of a variable impedance, which gives a direct reading of the reflection coefficient, this measurement can be readily performed with a form of bridge circuit. The bridge is balanced by making the reflection coefficient of the variable impedance equal to that of the unknown impedance.

The device employed here as the bridge is a hybrid T, a schematic representation of which is shown in fig. 13a. A wave entering the arm 1 divides equally at the branching point into two components, which enter the symmetrical arms 2 and 3. There is no direct transfer of energy from arm 1 to arm 4. Arm 2 is terminated by the unknown impedance Z_2 , and arm 3 by the variable (known) impedance Z_3 . The waves reflected from Z_2 and Z_3

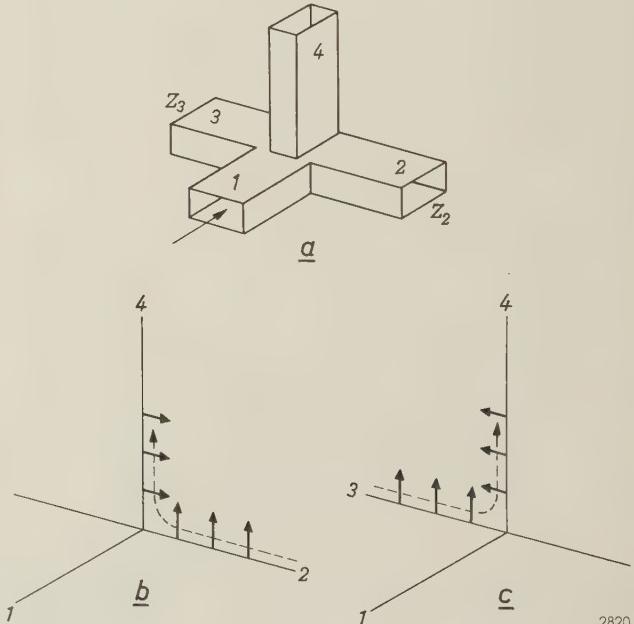


Fig. 13. a) Schematic representation of a hybrid Tee. A wave entering arm 1 divides into equal components which enter arms 2 and 3. These waves are reflected from the impedances Z_2 and Z_3 terminating these arms, and return partly into arm 1 (not considered here) and partly into arm 4. Here the waves cancel each other if Z_3 is equal in modulus and argument to Z_2 . b) The short arrows represent the electric field of the wave which is reflected from Z_2 and enters arm 4; the dashed arrow denotes the direction of propagation. c) The same for the wave reflected from Z_3 . When $Z_3 = Z_2$, the two waves cancel out in arm 4.

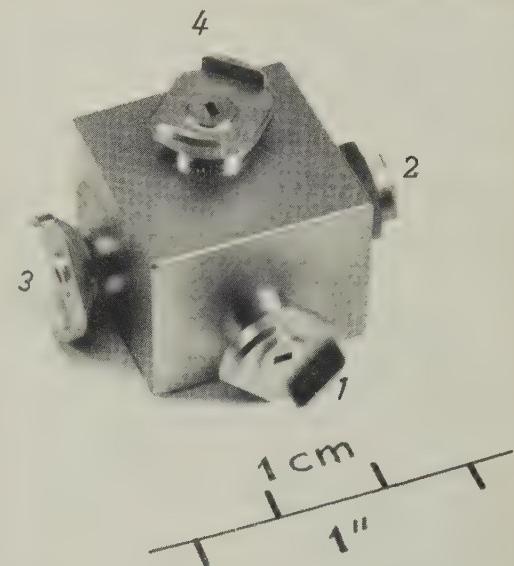
2820

return to the branching point and enter arm 4. When the modulus and the argument of Z_3 are made equal to those of Z_2 , the electrical field-strength of the wave reflected from Z_2 into arms 2 and 4 will be as shown in fig. 13b, while that for the wave reflected from Z_3 is as shown in fig. 13c. In arm 4 the field-strengths are equal and opposite, and the waves thus cancel each other. An indicator connected to a detector mounted in arm 4 will therefore give no deflection when the bridge is balanced, but does give a deflection as long as Z_3 is not equal to Z_2 .

Fig. 14 shows a hybrid T for 2 mm waves. A particularly important requirement for hybrid T's is that the common planes of symmetry of the Tee 1-2-3 and of the Tee 4-2-3 exactly coincide.

The principle of the variable impedance is illustrated in fig. 15¹⁵⁾. At the position of flange 1 the waveguide has a rectangular cross-section; via a transition 2, this becomes a circular cross-section 3. The latter contains two absorbing vanes, 4 and 5, which again consist of thin metallized slivers of mica. Vane 4 is fixed, perpendicular to the electric field E of the incident wave; vane 5 is mounted on a metal plunger 6, which can be both rotated and displaced axially. If the planes of 4 and 5 subtend an angle α , the component $E \sin \alpha$, parallel to vane 5, is absorbed by this vane; the component $E \cos \alpha$, perpendicular to 5, is reflected by the plunger 6. Now $E \cos \alpha$ has a component $E \cos \alpha \sin \alpha$, which is absorbed in vane 4, and a component $E \cos^2 \alpha$, which emerges from 1 as

¹⁵⁾ See also F. C. de Ronde, A simple component for impedance measurements at cm and mm waves: the direct-reading variable impedance, Communic. Congrès internat. Circuits et antennes hyperfréquences, Paris 1957, Part I (Suppl. Onde électr. **38**, No. 376 bis), pp. 294-295, 1958.



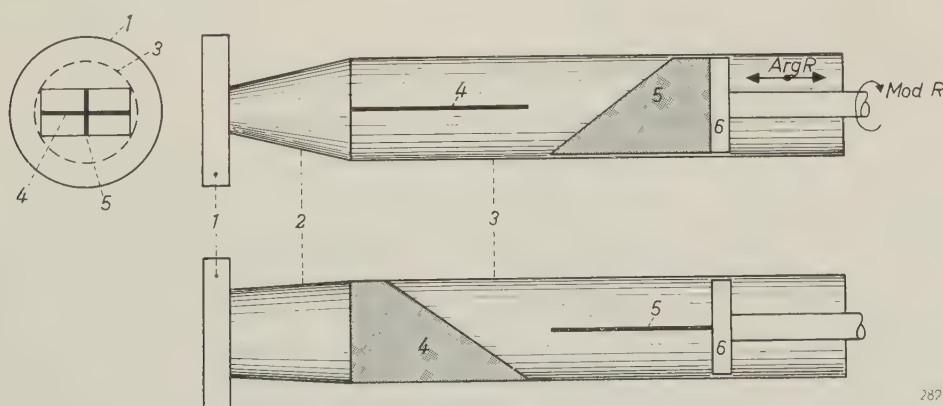
2809

Fig. 14. Hybrid T for 2 mm waves. Numbering of the arms as in fig. 13. The decoupling between arms 1 and 4 is greater than 30 dB.

a reflected wave, so that the modulus of the reflection coefficient is equal to $\cos^2 \alpha$. This modulus is varied by turning the vane 5 by means of the plunger spindle. The argument of the reflection coefficient is varied by displacing the reflecting surface, i.e. by moving the plunger axially, in or out. A displacement δl alters the phase φ of the reflected wave by

$$\delta\varphi = 2 \times \frac{\delta l}{\lambda_{gc}} \times 2\pi \text{ radians},$$

where λ_{gc} is the wavelength in the circular waveguide; the factor 2 is due to the fact that the wave



282

Fig. 15. Axial sections of the variable impedance along two mutually perpendicular planes, and end view. 1 flange with rectangular opening. 2 transition from rectangular to circular cross-section. 3 circular waveguide. 4 fixed vane. 5 vane capable of being rotated and axially displaced, fixed to plunger 6. When 5 is rotated, the modulus of the complex reflection coefficient is changed; when 5 is displaced axially, the argument is changed.

traverses the waveguide twice (incident and reflected waves).

As can be seen in fig. 16, the instrument gives a direct reading of both modulus and argument. The modulus scale, like the attenuation scale on the rotary attenuator, requires no previous calibration.



Fig. 16. Variable impedance as in fig. 15, for 2 mm waves. The modulus and argument of the reflection coefficient can be read from separate scales. If reflection losses are eliminated, the absolute error in the modulus reading is less than 0.05.

The rotary directional coupler

A directional coupler is a device that makes it possible to sample either the forward or backward wave in a waveguide, and thus to demonstrate, e.g., the presence of reflection. The most familiar type contains coupling holes between the main and a subsidiary waveguide. Its operation is illustrated in fig. 17. A wave travelling to the right through the main waveguide W_1 in fig. 17a produces no response in the detector mounted in the subsidiary waveguide W_2 ; the detector does respond, however, to a wave travelling in the opposite direction (fig. 17b). In this way one can detect the presence of a reflected wave.

A drawback of such directional couplers is that they function only in the frequency range where the (fixed) distance between the coupling holes is roughly equal to $\frac{1}{4}\lambda_g$. Instead of two large holes, it is the usual practice for various reasons to employ a series of small coupling holes, whose spacing must be extremely accurate; this involves considerable difficulties in manufacture. A third drawback is that the coupling is not variable. None of these disadvantages apply to the new directional coupler which we shall now describe.

The *rotary directional coupler*¹⁶⁾ (fig. 18a) consists of a circular waveguide W_1 (only part of which is shown in the figure, for simplicity) which terminates at both ends in a rectangular cross-section. The circular part, at end I, contains a metal strip Pol , the polarization strip, which can be regarded as lossless. At the left end (near I), this strip is exactly parallel with the broad faces of the waveguide, but it can be twisted into a helical surface by means of a rotatable section of the circular waveguide. The polarization strip terminates in a tapered absorption vane V , of the type already mentioned. This vane lies in exactly the same plane in which the polarization strip ends. Beyond the vane the circular waveguide contains a thin rod A , which acts as a coupling probe, and situated a quarter wavelength further on is a metal plate R , which acts as a reflector (see below). The probe and the reflector pass through the axis of the circular waveguide and are parallel to the broad sides of the rectangular ends. Transverse slots in the polarization strip and in the reflector prevent the occurrence of higher modes of oscillation.

The probe couples the circular waveguide with a rectangular subsidiary waveguide W_2 . To direct

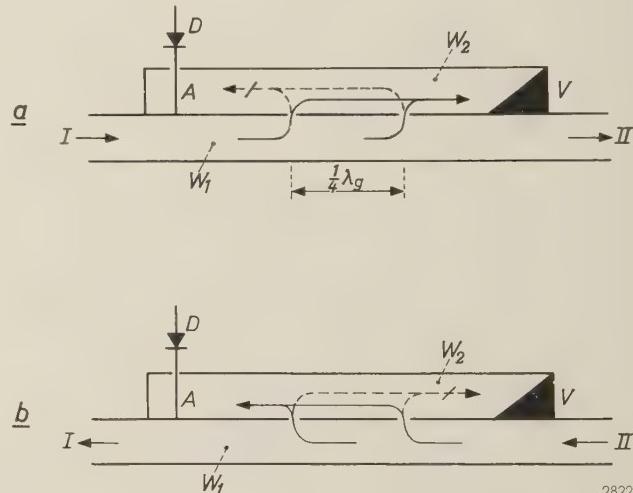


Fig. 17. Two-hole directional coupler. The common wall of the main waveguide W_1 and the subsidiary guide W_2 contains two holes one-quarter wavelength apart. A denotes a coupling probe, D a detector and V a dissipative vane or wedge.
 a) A part of a wave in W_1 travelling in the direction of the arrow passes through the coupling holes into W_2 , where it generates both backward and forward waves. The waves denoted by the solid arrow are in phase and thus reinforce one another, but they are absorbed by the vane V . The waves denoted by the dashed arrows differ in path-length by $2 \times \frac{1}{4}\lambda_g$, so that they cancel. No current is therefore induced in the probe.
 b) A wave travelling in the opposite direction in W_1 again produces a wave in W_2 which travels in the same direction as in W_1 , and now induces a current in the probe.

¹⁶⁾ Belgian patent application No. 464 685 of 16th December 1959.

all energy in this waveguide to one end — where there is usually a detector — the other end of W_2 is terminated by an adjustable plunger P .

The rotary directional coupler works as follows. We consider first a wave entering the coupler at I . The electric field E is thus perpendicular to the beginning of the polarization strip (fig. 18b). The wave then passes along this strip, which is twisted through an angle α . The direction of polarization

The reflector R ensures that the component $E \sin \alpha$ is coupled as fully as possible to the waveguide W_2 . For this purpose, R is situated a quarter wavelength behind the probe, and consists of a vertical metal plate, its function being to reflect vertically polarized waves back to the probe and let horizontally polarized waves pass through. The joint effect of the reflector R and the plunger P is to couple the component $E \sin \alpha$ almost entirely to W_2 .

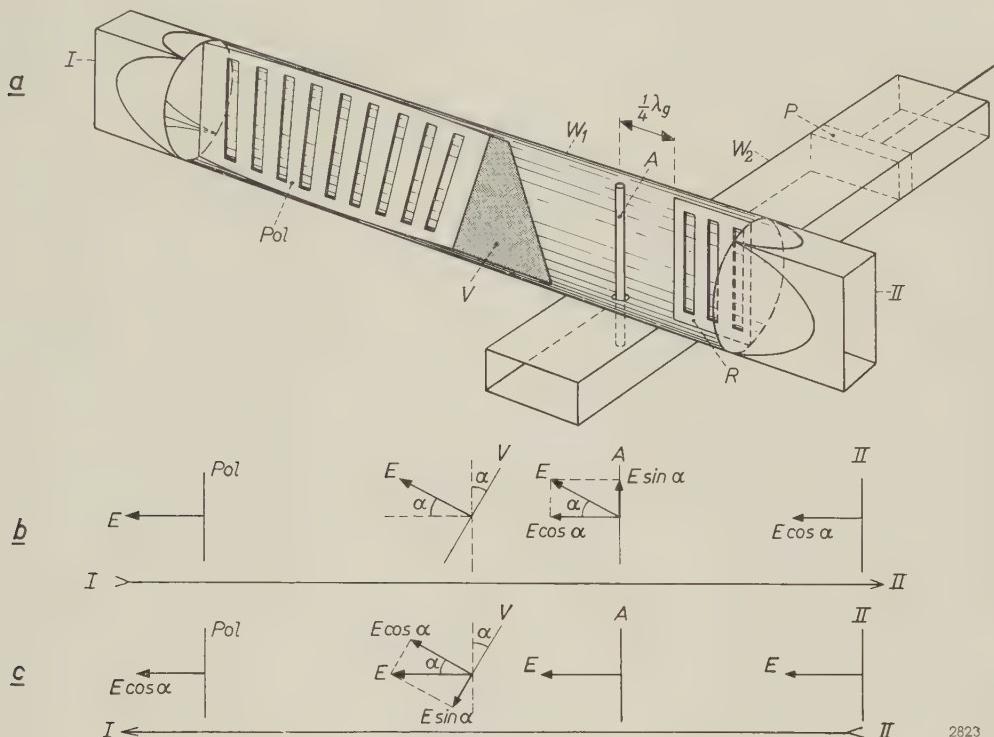


Fig. 18. a) Rotary directional coupler (schematic). W_1 circular waveguide terminating at both ends in rectangular openings I and II . Pol polarization strip which can be twisted from 0 to 45° by a central rotatable section of the waveguide wall. V vane. A coupling probe. R reflector. W_2 subsidiary waveguide with rectangular cross-section. P movable plunger. A wave passing from I to II is partly transferred to W_2 , the coupling being continuously variable and independent of frequency. No energy is coupled to W_2 in the case of a wave passing from II to I .

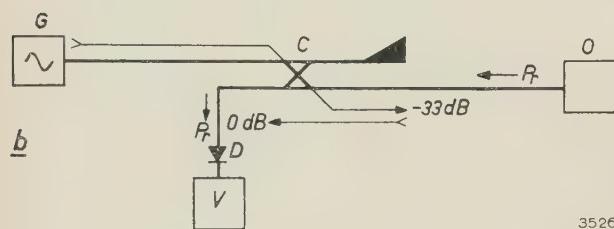
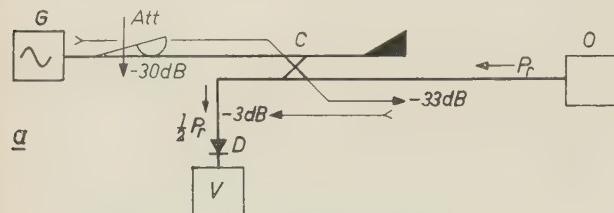
b) Electric field of a wave travelling from I to II , at the beginning of the polarization strip Pol , at the absorption vane V , at the probe A and in the opening II . The angle of twist is α .

c) The same for a wave from II to I ; from right to left; the field at II , at A , at V , and at the end of Pol .

thereby undergoes a rotation α , since the vector E remains perpendicular to the lossless strip. At the end of the strip, then, the field is unattenuated and is at right angles to the absorption vane, so that no energy is absorbed by the latter. Beyond the strip the field possesses a vertical component $E \sin \alpha$ and a horizontal component $E \cos \alpha$. The component $E \sin \alpha$ is coupled, via the probe, with the waveguide W_2 , whilst the component $E \cos \alpha$ passes on to the rectangular opening II .

A wave entering II in the opposite direction, with E again perpendicular to the broad sides of the rectangle (fig. 18c, right), passes the reflector, induces no current in the probe (so that energy in this direction is not transferred to W_2) and arrives at the absorption vane at a certain angle. The component $E \sin \alpha$ is absorbed by the vane, and the component $E \cos \alpha$ perpendicular to the vane reaches the twisted polarization strip, after which, still perpendicular to the strip and unattenuated,

it is turned back through an angle α to the position perpendicular to the broad sides of the rectangular guide, in which position it emerges from the rectangular opening I .



3526

Fig. 19. a) Illustrating the principle of a gyromagnetic resonance measurement. G microwave generator. Att variable attenuator. C fixed directional coupler. O specimen (solid in resonant cavity, in variable magnetostatic field). D detector. V millivoltmeter.

The directional coupler gives a coupling of, e.g., -3 dB in the path from G to O . To avoid saturating O , an extra attenuation of 30 dB is introduced, bringing the total attenuation to 33 dB . Of the total reflected power P_r , only half enters the detector via C .

b) The same layout, now with a rotary coupler. The required attenuation of 33 dB in example (a) can now be produced without a variable attenuator by adjusting the coupling from G to O to -33 dB . The coupling from O to D is then strong enough for almost the entire reflected power P_r to enter the detector.

The result is thus as follows.

- When waves travel in the direction from I to II , the component $E \sin \alpha$ is transferred to W_2 , whilst the component $E \cos \alpha$ passes through.
- When waves travel in the direction from II to I , no energy is transferred to W_2 , the component $E \sin \alpha$ is absorbed and the component $E \cos \alpha$ passes through.

The factor that indicates what fraction of the field of the waves travelling from I to II is transferred to the subsidiary waveguide is referred to as the coupling. The power

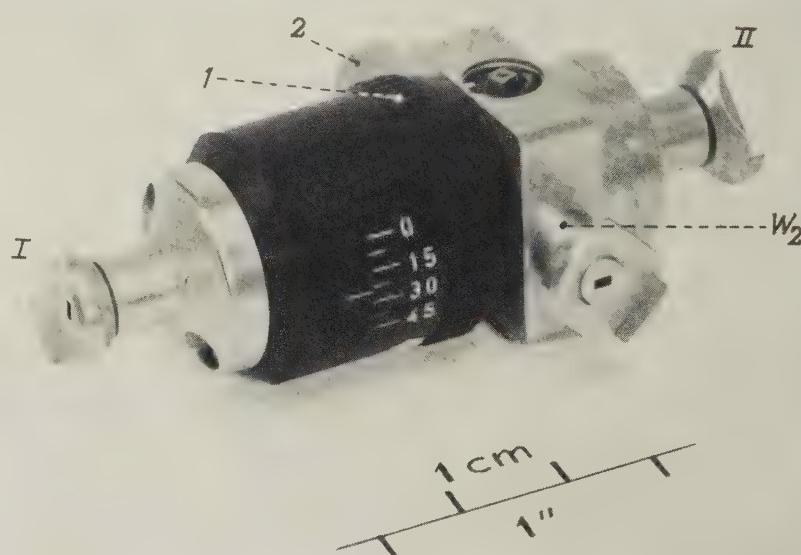
coupling in this case is thus $\sin^2 \alpha$ or, in decibels, $10 \log \sin^2 \alpha \text{ dB}$. The coupling effected by the rotary directional coupler is therefore continuously variable and moreover solely a function of the angle α , irrespective of the frequency. The value of the coupling can be read from a scale. In all other existing directional couplers the coupling is not variable and is to some extent frequency-dependent.

The usefulness of a continuously variable coupling may be illustrated by an example, in this case measurements of paramagnetic resonance in solids¹⁷⁾. (Such measurements are ordinarily performed in the centimetre wave region, but this does not detract from the advantages of a variable coupler over a fixed type.)

The principle of the measurement is illustrated in fig. 19a. Microwave energy from a generator G is conducted through a directional coupler C to the sample under investigation O , which is contained in a resonant cavity, itself situated in a variable magnetostatic field. Through the same directional coupler a part of the reflected power P_r is transferred to the detector D . When the magnetostatic field is varied, resonance is observed as a sharp dip in the reflected power. The sensitivity is optimum when the coupling is -3 dB . This value is obtained with a hybrid T (preferably a magic T, i.e. a hybrid T free from reflection). For this reason, and also because of its good directivity (this term is explained below), the magic T is often used as a directional coupler in this kind of measurement.

A coupling of -3 dB , however, may give rise to saturation

¹⁷⁾ See e.g. J. S. van Wieringen, Paramagnetic resonance, Philips tech. Rev. **19**, 301-313, 1957/58.



2811

Fig. 20. Rotary directional coupler for 2 mm waves. I , II and W_2 as in fig. 18a. 1 rotatable ring for twisting the polarization strip; the angle of twist α is read from the scale. 2 knob for adjusting the plunger in W_2 . Directivity $>25\text{ dB}$, standing-wave ratio <1.16 .

in the specimen O , that is to say the incident power is so high that the reflected power is virtually independent of it. To remain below the saturation region one must therefore reduce the incident power. The most obvious method is to replace the fixed directional coupler by one which gives a weaker coupling, but this is cumbersome and entails stepwise regulation. Another method is to insert a variable attenuator (*Att.*, fig. 19a) between the generator and the directional coupler. Suppose that the power must be attenuated by 30 dB to remain below the saturation limit. The attenuation between the generator and the specimen is then 33 dB (fig. 19a). This can be achieved more simply by substituting for the attenuator and the fixed 3 dB directional coupler a variable rotary coupler adjusted to -33 dB (fig. 19b). This dispenses with the variable attenuator. A further advantage is that, with a weak coupling from G to O , the coupling from O to D is strong enough to cause virtually all the reflected power P_r to arrive in D . Where a fixed directional coupler of -3 dB is used (fig. 19a), half of P_r is lost.

In fig. 19b the rotary coupler is connected as follows (cf. fig. 18a): the generator at the free end of the subsidiary waveguide W_2 , the object at side I and the detector at side II .

For the rotary coupler it is found to be sufficient if α is variable from 0 to 45° . Theoretically, the coupling is then variable from $-\infty$ dB to -3 dB. Owing to mechanical imperfections, the weakest coupling in practice is between -30 and -40 dB.

For the same reason W_2 still receives a small fraction of the energy of a wave passing through the rotary coupler from II to I . This imperfect directional effect — found in all such devices — is expressed quantitatively in the directivity. This is the ratio of the power coupled to the subsidiary waveguide as a result of a wave from I to II , to that of an equally

strong wave from II to I . The directivity of the rotary coupler is thus better the tighter the coupling.

A photograph of the 2 mm rotary directional coupler can be seen in fig. 20.

Some measuring arrangements using the components described in the foregoing will be the subject of Part II of this article. This will also provide an opportunity to describe various other components not dealt with here.

Summary. For measurements in the 2 mm wavelength region a comprehensive range of waveguide components has been developed, the principal of which are reviewed in this article. The 2 mm waves are produced by using a silicon diode to double the frequency of the 4 mm waves generated by a reflex klystron, type DX 151. The detector used is also a silicon diode. The matching is effected by means of waveguide plungers and a pivoting screw tuner (a variant of the sliding screw tuner). Use is made either of synchronous detection or of a selective indicator, the 2 mm waves being modulated in amplitude, e.g. by superposing an audio-frequency voltage of about 1 V on the frequency multiplier. Without a frequency multiplier the PIN modulator may be used, which is an electrically controlled attenuator consisting of a germanium wafer having a P , an I and an N region. The waveguides are generally rectangular in cross-section (0.83×1.66 mm, or $0.0325'' \times 0.065''$) and gold-plated inside and out. They are joined together by claw flanges, which allow particularly compact assemblies. Three components are described in some detail: the rotary attenuator, the variable impedance and the rotary directional coupler. The properties of these components are entirely or largely independent of frequency. The rotary directional coupler gives a coupling which is continuously variable from -3 to -30 or -40 dB, and can be read from a dial. The variable impedance is used in conjunction with a hybrid T and gives a direct reading of the modulus and argument of the reflection coefficient.

X-RAY DETERMINATION OF CRYSTAL STRUCTURES

by P. B. BRAUN and A. J. van BOMMEL.

548.735

Since the fundamental work of Von Laue and Bragg (1912), X-ray diffraction analysis has grown into an extremely important technique for studying the solid state. It has led to an impressive series of structure determinations, including such spectacular ones as vitamin B 12, strychnine and penicillin.

The article below deals with the principles underlying X-ray diffraction, discussing the relation between the structure and the diffraction patterns, and how the one can be derived from the other. By way of illustration, examples of structures determined in Eindhoven are discussed, special mention being made of work on a substance from a class of magnetic materials known as ferroxplana.

Crystals of differing structure also differ in the way they scatter X-rays. This fact underlies a method of identifying crystalline substances, by means of the characteristic diffraction patterns to which these substances give rise when irradiated with X-rays¹⁾. (Diffraction patterns have therefore been described as the finger prints of crystals.) Diffraction patterns can moreover be used for determining the crystal structures of substances. The

detailed analysis of the crystalline and molecular structure of penicillin is one of the remarkable results achieved in this way. As appears in fig. 1, the structure of the penicillin molecule is very complicated, and its elucidation was therefore a landmark in the development of X-ray analysis. Other outstanding results in this field have been the determinations of the structure of strychnine and of vitamin B 12²⁾.

At the same time X-ray diffraction studies can lead to a better understanding of those physical and chemical properties that are related to the crystal structure of solids and molecules. Structure analysis is accordingly indis-

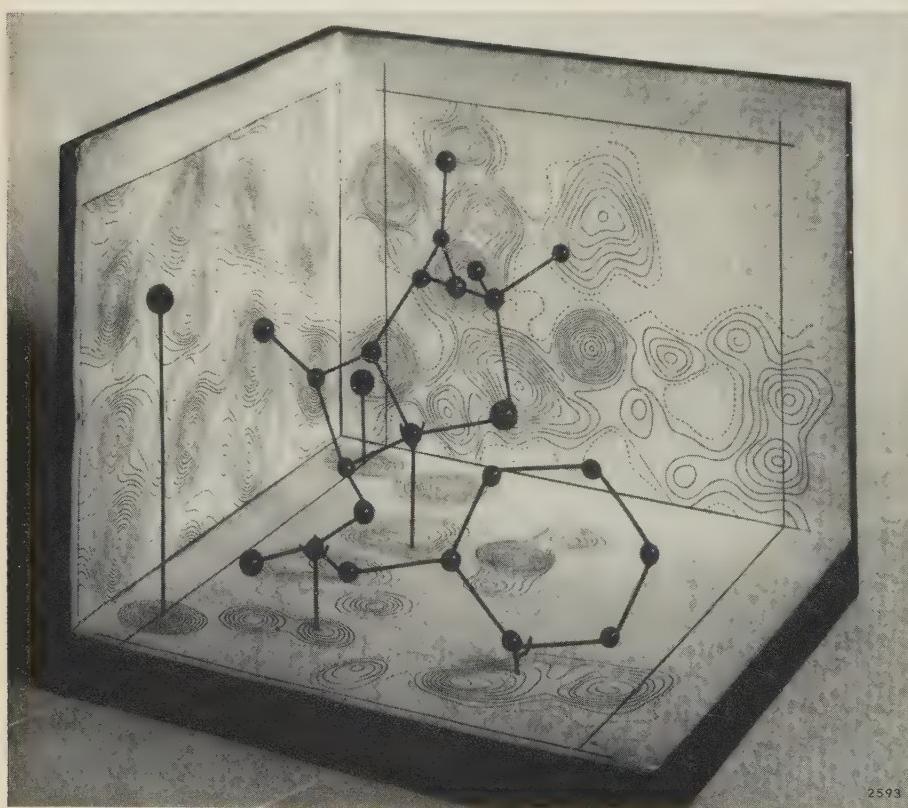


Fig. 1. Model of the benzyl penicillin molecule, obtained from the "Fourier maps" also shown. The latter represent three projections of the molecule, the "contours" being lines of constant electron density, and the "peaks" corresponding to the projections of atoms. These Fourier maps were constructed from data obtained from X-ray diffraction patterns. (From: E. Chain, Endeavour No. 28, Oct. 1948, p. 152.)

¹⁾ See e.g. W. G. Burgers, Philips tech. Rev. **5**, 157, 1940; W. Parrish and E. Cisney, Philips tech. Rev. **10**, 157, 1948/49.

²⁾ D. Crowfoot, C. W. Bunn, B. W. Rogers-Low and A. Turner-Jones, The chemistry of penicillin, Princeton University Press, 1949. J. H. Robertson and C. A. Beevers, The crystal structure of strychnine hydrogen bromide, Acta cryst. **4**, 270-275, 1951. C. Bokhoven, J.C. Schoone and J. M. Bijvoet, The Fourier synthesis of the crystal structure of strychnine sulphate pentahydrate, Acta cryst. **4**, 275-280, 1951. D. Crowfoot Hodgkin et al., The structure of vitamin B 12, Proc. Roy. Soc. A **242**, 228-263, 1957.

pensable in the study of ceramic magnetic materials, such as ferroxdure and ferroxplana³⁾.

To understand the principles on which X-ray structure analysis is based, it is necessary first of all to know what information is contained in an X-ray diffraction pattern, and how the structure of a substance can be deduced from such a pattern. This will be the subject of the first half of this article⁴⁾. The second half deals briefly with three examples taken from researches carried out in recent years at the Philips Research Laboratories in Eindhoven.

Electron and neutron diffraction analysis are analogous to X-ray diffraction analysis. These methods are all based on the same principles, but differ in technique. They also differ to some extent in the information they yield, so that they supplement one another very usefully. An article on neutron diffraction will appear in these pages in the near future; the various differences from X-ray diffraction will be further discussed there. The examples of structure determinations given in both articles will relate to the same substances, so that it will be readily possible to compare the two methods.

One final prefatory remark: X-rays are scattered solely by electrons. This means that the data derived from X-ray diffraction patterns can only relate to

³⁾ See e.g. J. J. Went, G. W. Rathenau, E. W. Gorter and G. W. van Oosterhout, Philips tech. Rev. **13**, 194, 1951/52; G. H. Jonker, H. P. J. Wijn and P. B. Braun, Philips tech. Rev. **18**, 145, 1956/57.

⁴⁾ Details of this subject will be found in such works as: J. M. Bijvoet, N. H. Kolkmeijer and C. H. MacGillavry, Röntgenanalyse van kristallen, Centen, Amsterdam 1948; H. Lipson and W. Cochran, The determination of crystal structures, Bell, London 1953; R. W. James, The crystalline state, Vol. II. The optical principles of the diffraction of X-rays, Bell, London 1948; J. Bouman, Theoretical principles of structural research by X-rays, Handbuch der Physik, Vol. **32**, 95-237, Springer, Berlin 1957. For the techniques used for recording diffraction patterns, see e.g. M. J. Buerger, X-ray crystallography, Wiley, New York 1942; H. P. Klug and L. E. Alexander, X-ray diffraction procedures, Wiley, New York 1954.

these particles. To determine a crystalline structure it is necessary to find the sites of the atoms, but this can be done by localizing the electron clouds. We can thus formulate the direct objective of the X-ray analysis of crystal structure as *the determination of the spatial distribution of the electron density in a crystal*.

A useful way of describing the electron density in a crystal

It is characteristic of a crystal that it possesses three-dimensional periodicity. This means that it is possible to choose a volume element, by the regular stacking of which the whole crystal can be built up. The smallest volume element with which this can be done is called the *unit cell*.

The three-dimensional periodicity is, of course, shared by the *electron density*.

Now a periodic function having a period d can always be resolved into a series of sine waves whose periods are $d, d/2, d/3$, etc. The process by which a function is resolved into these components is called *Fourier analysis*, whilst the building-up of a function from these components is referred to as *Fourier synthesis*.

In view of its periodicity, the electron density in a crystal can also be resolved into Fourier components. These constitute a collection of plane stationary waves having different directions. All the waves having any one direction comprise a set having wavelengths which are sub-multiples of the periodicity d in that direction. The Fourier components are referred to here as *density waves* or *density components* (see figs. 2 and 3).

The electron density may thus be expressed as follows:

$$\varrho(x, y, z) = \sum_{h}^{\infty} \sum_{k}^{\infty} \sum_{l}^{\infty} A_{hkl} e^{2\pi i \left(\frac{hx}{a} + \frac{ky}{b} + \frac{lz}{c} \right)}. \quad (1)$$

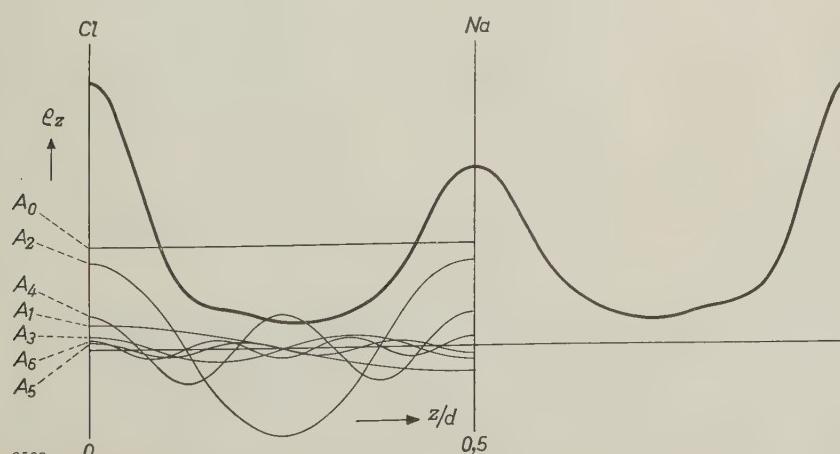


Fig. 2. The thick upper curve is obtained by adding together the lower series of cosine waves (A_0, A_1, A_2 etc.) whose wavelengths are all integral sub-multiples of d . The thick curve represents the electron density (ϱ_z) between the octahedral planes of rock-salt, while A_0, A_1, A_2 , etc. are the amplitudes of the Fourier components of this curve, referred to here as "density waves". z is the direction through the crystal (viz. normal to the octahedral planes), d the lattice spacing in that direction.

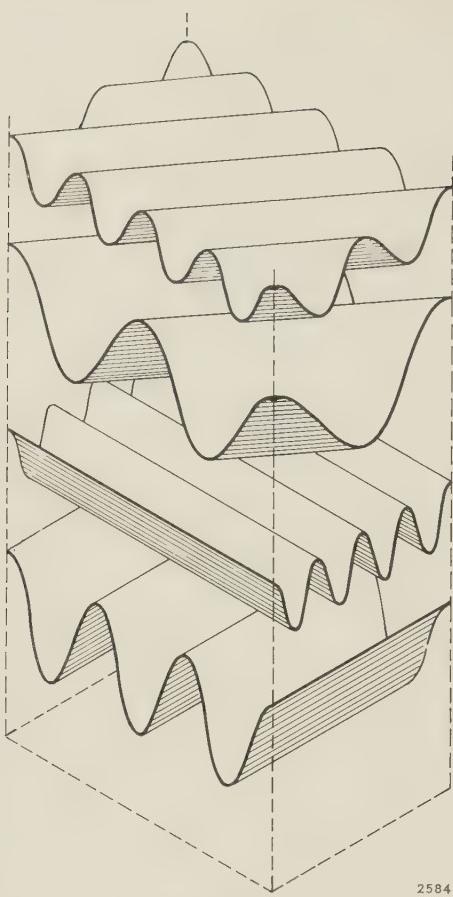


Fig. 3. Schematic representation of various density waves, in which the electron density is depicted by the geometrical amplitude. Two of the waves shown have the same direction, their wavelengths being in the ratio 1 : 2; the two other waves shown have different directions. To find the electron density distribution in a crystal, a number of such waves must be added together.

The term on the left is the electron density as a function of the coordinates x , y and z . The terms on the right constitute the corresponding electron density waves. The wavelength and direction of each wave are expressed, with the aid of the integers h , k and l , in terms of the longest repetition distances a , b and c of three waves whose directions do not lie in one plane⁵⁾. A_{hkl} is a complex quantity whose modulus and argument indicate the amplitude and phase, respectively, of the relevant density wave.

The volume of the parallelepiped whose edges are a , b and c will generally be taken as small as possible; in this way we thus arrive at the above-mentioned unit cell.

⁵⁾ In the text books quoted the symbols h , k and l are given the meaning of indices of lattice planes. In a brief discussion of the principles of structural analysis, this is neither necessary nor desirable. No mention will therefore be made of lattice planes in this article, and the Bragg law will accordingly appear in a somewhat modified form.

Sometimes to bring out the symmetry of the crystal, a unit cell is adopted that is an integral number of times larger than the smallest possible. In that case the quantity A_{hkl} in formula (1) corresponding to given combinations of h , k and l is reduced to zero (a case of general extinctions).

We shall now show that the electron density waves can be deduced directly from diffraction patterns. We shall consider the simplest conceivable case, in which the electron density may be described in terms of a single density wave (this is never so in reality), and show what the form of the diffraction pattern is in that case. We shall then do the same for an electron distribution consisting of a number of density waves, as found in an actual crystal.

Diffraction by electron density waves

The question in what directions a single density wave scatters is answered most easily if we consider the effect of successive "cross-sectional layers" of the density wave, as depicted in fig. 4. Each of these layers acts as a kind of mirror. The scattered radiation is thus to be sought in the direction corresponding to simple reflection from these layers.

In general, the total yield in reflected radiation from all layers is zero. A perceptible reflection occurs only under the condition that the angle of incidence Θ satisfies the relation:

$$2d \sin \Theta = \lambda, \dots \quad (2)$$

where λ is the wavelength of the radiation and d the wavelength of the density wave. In that case the amplitude of the reflected waves has a value proportional to the amplitude of the density wave.

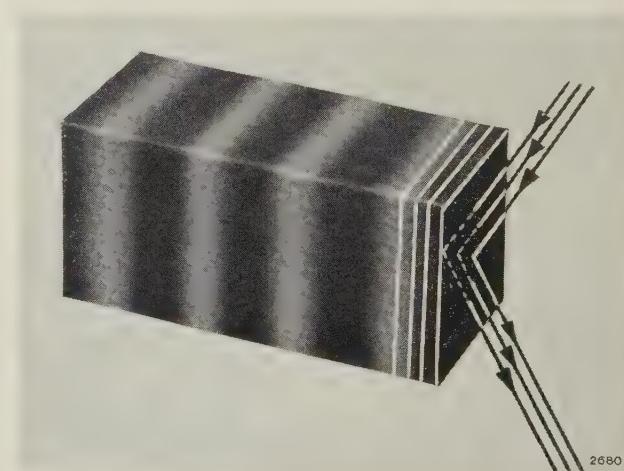


Fig. 4. Representation of a density wave with an X-ray beam incident upon it. The variation of the electron density in the wave is indicated by the shading. On the right, three "layers" of the density wave are represented, each of which acts as a reflector for X-rays. Rays reflected from the various layers interfere with one another and cause, in general, extinction, except at one specific angle of incidence, where a reflection of finite intensity occurs.

These two relations, concerning the angle and the amplitude of reflections, are the foundations of the X-ray analysis of crystal structures. They show that a reflection provides information on no less than three properties of a density wave, viz. the amplitude, the wavelength and the direction. The intensity of the reflection is a measure of the amplitude of the density wave. Given the wavelength of the radiation, we can calculate the wavelength of the density wave from the angle of reflection. Thirdly the position of the crystal together with the direction and angle of the reflection determines the direction of the density wave.

The above can easily be demonstrated vectorially with the aid of fig. 5.

The difference in path length of the rays drawn in fig. 5 is equal to the component of the vector \mathbf{r} along \mathbf{S} minus the component of \mathbf{r} along \mathbf{S}_0 , or:

$$\mathbf{r} \cdot (\mathbf{S} - \mathbf{S}_0).$$

The corresponding phase difference is:

$$\frac{2\pi \mathbf{r} \cdot (\mathbf{S} - \mathbf{S}_0)}{\lambda}.$$

The total amplitude F of the radiation, caused by the scattering volume, is obtained by adding together the contributions from all volume elements dV , each having an electron density $\varrho(\mathbf{r})$, and taking the mutual phase differences into account:

$$F = \int_V \varrho(\mathbf{r}) e^{-2\pi i \frac{\mathbf{r} \cdot (\mathbf{S} - \mathbf{S}_0)}{\lambda}} dV. \quad \dots \quad (3)$$

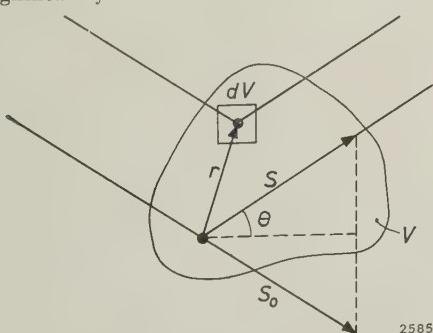
Regarding the scattering body as an electron density wave, we can write for $\varrho(\mathbf{r})$:

$$\varrho(\mathbf{r}) = A e^{2\pi i \frac{\mathbf{r} \cdot \mathbf{z}}{d}}, \quad \dots \quad (4)$$

where \mathbf{z} is the unit vector indicating the direction of the density wave, and A and d represent the amplitude and the wavelength, respectively. Using (4), the expression (3) becomes:

$$F = \int_V A e^{2\pi i \frac{\mathbf{r} \left(\frac{\mathbf{z} - \mathbf{S} - \mathbf{S}_0}{d} \right)}{\lambda}} dV. \quad \dots \quad (5)$$

Only when the exponent is zero will this integral have a value that is significantly different from zero.



2585

Fig. 5. Determination of the reflection condition for a density wave. \mathbf{S}_0 and \mathbf{S} are unit vectors denoting respectively the directions of incidence and scattering of an X-ray beam, two rays of which are shown. \mathbf{r} is the vectorial distance between two scattering points, dV is a volume element of the total volume V , and Θ is the semi-angle of reflection.

This gives us the condition for reflection:

$$\frac{\mathbf{z}}{d} = \frac{\mathbf{S} - \mathbf{S}_0}{\lambda}.$$

This means:

- a) The two vectors $(\mathbf{S} - \mathbf{S}_0)$ and \mathbf{z} must have the same direction, i.e. the directions of incidence (\mathbf{S}_0) and reflection (\mathbf{S}) lie with \mathbf{z} in one plane and make the same angle with \mathbf{z} . In other words, the layers into which we have divided the density wave act as reflectors.
- b) The magnitudes of the vectors must be equal. Since $|\mathbf{S} - \mathbf{S}_0| = 2 \sin \Theta$ (see fig. 5), $1/d = 2 \sin \Theta/\lambda$ or $2d \sin \Theta = \lambda$. These conditions being fulfilled, the integral gives:

$$F = VA, \quad \dots \quad (6)$$

i.e. the amplitude of the reflected ray is proportional to the amplitude of the electron density wave considered ⁶⁾.

We have seen that the electron density in a crystal can be described in terms of a group of density waves. Each such wave, independent of the other waves, is capable of "reflecting" in the manner discussed. A further point of importance is that each reflection can generally be intercepted separately from the other reflections. (Since no two waves are identical in direction as well as wavelength, they all "reflect" under different conditions.) This makes it possible to derive from each density wave the information present in its reflection.

To obtain this information it is convenient to use various different ways of irradiating the crystal(s). Weissenberg patterns are particularly valuable in structural analysis, but Laue and Debye-Scherrer patterns and rotation photographs are also often useful for this purpose (see fig. 6). The reflections may be recorded photographically or they may be measured with the aid of Geiger counters, proportional counters or scintillation counters ⁷⁾.

The underlying principle of structure analysis

Summarizing the foregoing, the principles of the X-ray determination of crystal structures may be expressed as follows.

The electron density in a crystal can be described as the sum of a number of Fourier components. Each

⁶⁾ In the derivation given here we have assumed that a ray, once reflected, leaves the crystal without again being reflected. The so-called "dynamic theory", which takes account of multiple reflections, produces slightly different results from the "kinematic theory" used here. The latter theory proves to be adequate for most crystals in practice, but corrections may sometimes be called for (extinction corrections). The situation is different where perfect crystals with no defects are concerned. In their case "dynamic" effects are very important. See e.g. L. P. Hunter, Anomalous transmission of X-rays by single crystal germanium, Proc. Kon. Ned. Akad. Wet. **B** 61, 214-219, 1958.

⁷⁾ See e.g. W. Parrish, E. A. Hamacher and K. Lowitzsch, The "Norelco" X-ray diffractometer, Philips tech. Rev. **16**, 123-133, 1954/55; P. H. Dowling, C. F. Hendee, T. R. Kohler and W. Parrish, Counters for X-ray analysis, Philips tech. Rev. **18**, 262-275, 1956/57.

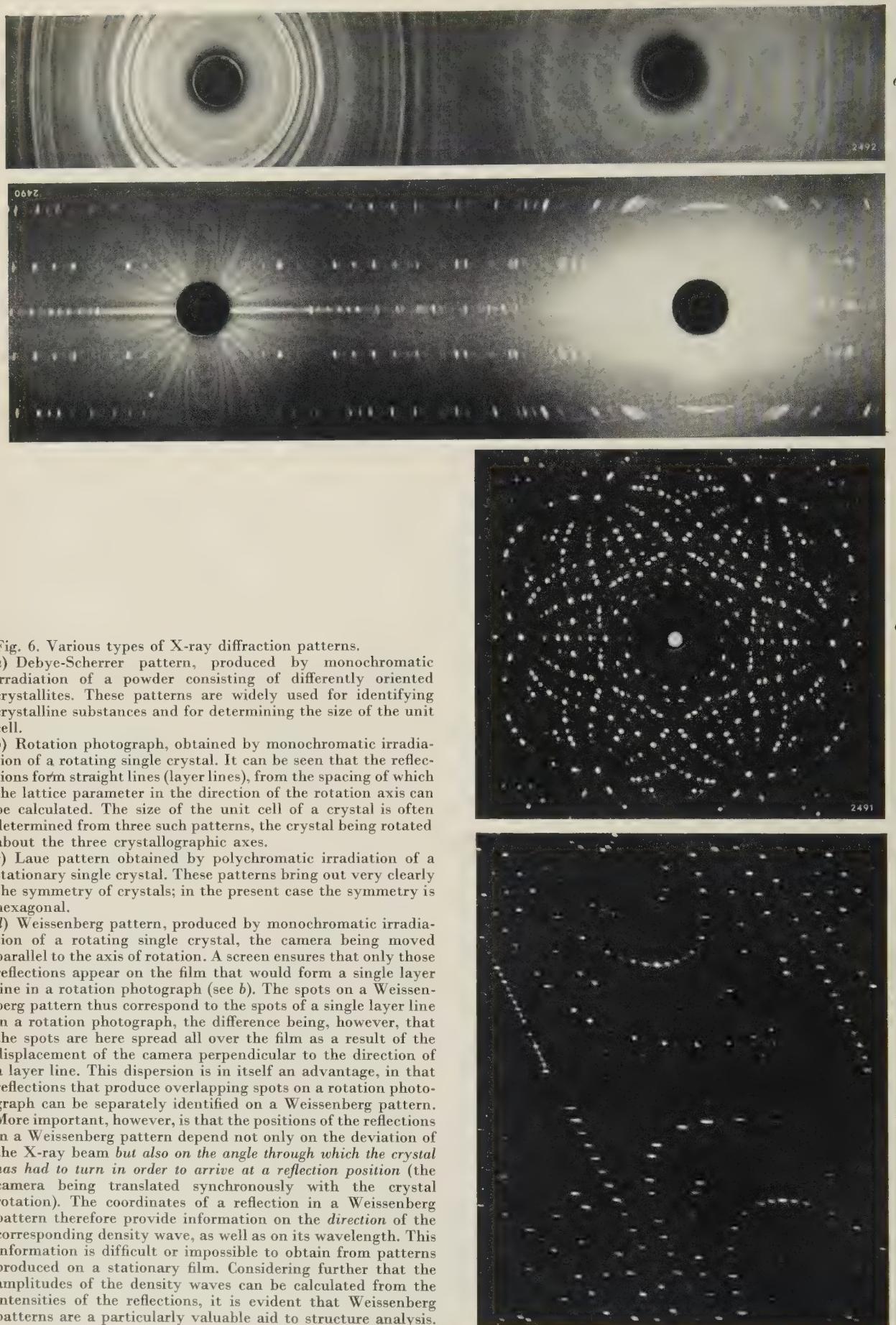


Fig. 6. Various types of X-ray diffraction patterns.

a) Debye-Scherrer pattern, produced by monochromatic irradiation of a powder consisting of differently oriented crystallites. These patterns are widely used for identifying crystalline substances and for determining the size of the unit cell.

b) Rotation photograph, obtained by monochromatic irradiation of a rotating single crystal. It can be seen that the reflections form straight lines (layer lines), from the spacing of which the lattice parameter in the direction of the rotation axis can be calculated. The size of the unit cell of a crystal is often determined from three such patterns, the crystal being rotated about the three crystallographic axes.

c) Laue pattern obtained by polychromatic irradiation of a stationary single crystal. These patterns bring out very clearly the symmetry of crystals; in the present case the symmetry is hexagonal.

d) Weissenberg pattern, produced by monochromatic irradiation of a rotating single crystal, the camera being moved parallel to the axis of rotation. A screen ensures that only those reflections appear on the film that would form a single layer line in a rotation photograph (see b). The spots on a Weissenberg pattern thus correspond to the spots of a single layer line in a rotation photograph, the difference being, however, that the spots are here spread all over the film as a result of the displacement of the camera perpendicular to the direction of a layer line. This dispersion is in itself an advantage, in that reflections that produce overlapping spots on a rotation photograph can be separately identified on a Weissenberg pattern. More important, however, is that the positions of the reflections in a Weissenberg pattern depend not only on the deviation of the X-ray beam *but also on the angle through which the crystal has had to turn in order to arrive at a reflection position* (the camera being translated synchronously with the crystal rotation). The coordinates of a reflection in a Weissenberg pattern therefore provide information on the *direction* of the corresponding density wave, as well as on its wavelength. This information is difficult or impossible to obtain from patterns produced on a stationary film. Considering further that the amplitudes of the density waves can be calculated from the intensities of the reflections, it is evident that Weissenberg patterns are a particularly valuable aid to structure analysis.

density component gives rise to an X-ray reflection which provides information on the amplitude, wavelength and direction of that particular component. From all the reflections from a crystal we have a (virtually) complete picture of all the electron density components, and hence in principle of the electron distribution in a crystal. The determination of this electron distribution — and with it the structure of the crystal — amounts to carrying out a Fourier synthesis, i.e. adding all components of the Fourier series to yield the required electron distribution function. It should be noted that the synthesis involved is three-dimensional, and is therefore extraordinarily complicated. A common practice is therefore to carry out one-dimensional and two-dimensional Fourier syntheses, which produce projections of the structure (projections on a line and on a plane, respectively). From two or three such projections it is possible to deduce the spatial configuration of the atoms. (This is well brought out in fig. 1.)

The phase problem

However close the solution may now seem, there is still a lot to be done before we can determine the structure of a crystal from the summing of the density components. Although at first sight it might seem that a mere addition is called for, closer examination shows that one quantity is still missing. This brings us up against an obstacle briefly referred to as the "phase problem". Overcoming this obstacle often costs the crystallographer most of the time he spends on analysing the structure of a crystal.

To carry out the Fourier summation correctly we must know, apart from the amplitude, wavelength and direction of the constituent waves, the *phase* of the waves, which is given by the argument of A_{hkl} in equation (1). This is demonstrated in fig. 7.

Unfortunately, the phase of the density components cannot be observed; unlike the amplitude,

wavelength and direction, it cannot be deduced from the reflections. The intensity of a reflected ray is a measure of the amplitude of the corresponding density wave, but tells us nothing about its phase.

It is for this reason that the diffraction pattern of a crystal cannot be directly interpreted to give a pattern of the electron density in a crystal. Not until the phases have been determined, by the methods about to be discussed, is it possible to carry out the superposition of Fourier components and so produce the required picture of the electron clouds which show the arrangement of the atoms or molecules in a crystal.

Phase determination

Just as knowledge of the phases of the reflections found by experiment would enable us to solve the structure of a crystal, it is possible conversely to calculate the phases from a known structure. This implies a possible method of solving the phase problem, for if we can produce a model representing a reasonable approximation of the true structure, the phases of the strongest reflections can often be determined with a fair degree of accuracy. These data are then used for carrying out a (provisional) Fourier synthesis. Although the picture obtained in this way is usually very rough, it often leads to a better approximation to the actual structure than the original model. With this improved approximation, better calculation of the phases can be made, and so on. In this procedure, phase calculation alternates with Fourier synthesis, thus producing step by step an increasingly more accurate picture of the electron density in the crystal.

The question now is how to arrive at an initial model to start the above procedure. The search for a suitable model leads along well-trodden paths, each attempt requiring from the investigator considerable crystallographic experience, deductive ability and intuition. For each attempt we have to

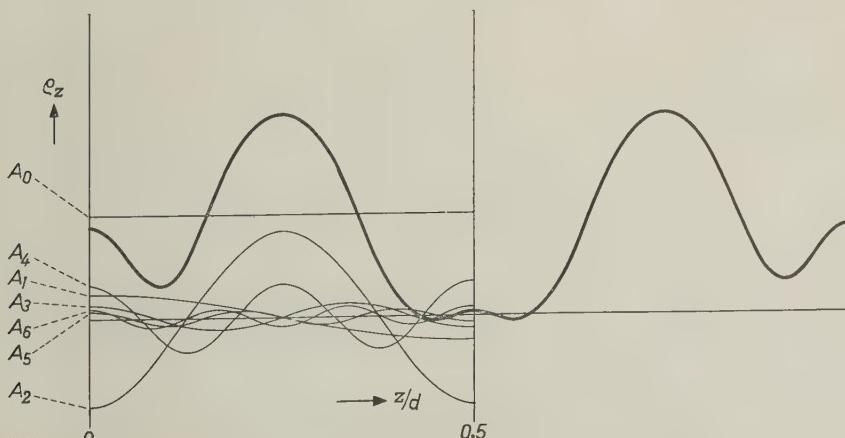


Fig. 7. Result of a Fourier synthesis, done with the same Fourier components as in fig. 2, but with the component A_2 having a different phase. The result is entirely different from the density distribution obtained in fig. 2.

ask the question: what is known of the crystal under investigation, apart from the X-ray data, which can help us to make a provisional guess at the structure?

In the first place the constituent atoms of a crystal are usually known from chemical analysis. The scattering contribution of each atom ("atomic structure factor") has been calculated for all elements and these are published in tables. We can also measure the density of crystals and calculate the number of atoms contained per unit volume. For our purposes it is particularly interesting to know the number of atoms contained in the unit cell, since the dimensions of this cell can be derived from the diffraction pattern. Having ascertained the number of atoms of each kind contained in this volume of known dimensions, it remains to find the spatial arrangement of the atoms.

Obviously, only those arrangements will be considered that agree with the properties of the crystal (chemical, optical, magnetic, morphological and particularly the symmetry properties). Taking these data into account, many possibilities can be eliminated, but it is hardly ever possible to reduce the remaining possible structures to one. Nevertheless, if only a few models are left after this elimination, it is now usually a practicable proposition to try each of them in turn. This is referred to simply as the "trial and error" method.

Knowledge of the *interatomic distances* can be of great help in the construction of a model. This knowledge is useful in the process of elimination described, and is sometimes decisive for the success of the analysis.

Information on the distances (magnitude and direction) between the atoms in a crystal (but not the sites of the atoms) can be obtained by Patterson's method. This is based on the dependence of the intensities of the reflections on the interatomic distances. The method consists in carrying out a so-called Patterson synthesis, which is a Fourier synthesis in which the components to be summed are simply the reflection intensities. This can be done *without* the phases of the reflections being known.

To show that something can indeed be learnt about interatomic distances from the reflection intensities, we shall first write formula (1) in a simpler notation:

$$\varrho(\mathbf{r}) = \sum_{\mathbf{H}} A_{\mathbf{H}} e^{2\pi i \mathbf{H} \cdot \mathbf{r}}, \dots \quad (7)$$

where \mathbf{r} is the vector from the origin to the point x, y, z and $\mathbf{H} \cdot \mathbf{r}$ represents $h \frac{x}{a} + k \frac{y}{b} + l \frac{z}{c}$.

From this equation we shall now derive another in which the intensities occur. In this connection the following points should be noted: a) $A_{\mathbf{H}}$ is in general a complex quantity; b) the square of the modulus of this quantity, $|A_{\mathbf{H}}|^2$ or $A_{\mathbf{H}} A_{\mathbf{H}}^*$,

where * denotes the complex conjugate of $A_{\mathbf{H}}$, is related, from eq. (6), to the intensity $I_{\mathbf{H}}$ ($\propto FF^*$) of the corresponding reflection by:

$$A_{\mathbf{H}} A_{\mathbf{H}}^* = \frac{I_{\mathbf{H}}}{V^2}. \dots \quad (8)$$

Of course, eq. (7) is also valid when $A_{\mathbf{H}}$ is replaced by its complex conjugate. Moreover, the equation also remains valid when the independent variable \mathbf{r} is replaced by $\mathbf{r}-\mathbf{u}$, where \mathbf{u} is any arbitrary vector. (By this device, as will be shown, we introduce the interatomic distance \mathbf{u} .) We may thus write:

$$\varrho^*(\mathbf{r}-\mathbf{u}) = \sum_{\mathbf{H}} A_{\mathbf{H}}^* e^{-2\pi i \mathbf{H} \cdot (\mathbf{r}-\mathbf{u})}. \dots \quad (9)$$

Multiplication of (7) and (9), distinguishing the vectors \mathbf{H} in (7) and (9) by \mathbf{H} and \mathbf{H}' , gives:

$$\varrho(\mathbf{r}) \varrho^*(\mathbf{r}-\mathbf{u}) = \sum_{\mathbf{H}} \sum_{\mathbf{H}'} A_{\mathbf{H}} A_{\mathbf{H}'}^* e^{2\pi i \mathbf{H}' \cdot \mathbf{u}} e^{2\pi i (\mathbf{H}-\mathbf{H}') \cdot \mathbf{r}}. \dots \quad (10)$$

We shall henceforth omit the * sign in $\varrho^*(\mathbf{r}-\mathbf{u})$, since this quantity is real and thus identical to its complex conjugate. Integrating both sides over the whole volume of the unit cell gives:

$$\int_V \varrho(\mathbf{r}) \varrho(\mathbf{r}-\mathbf{u}) dV(\mathbf{r}) = \sum_{\mathbf{H}} \sum_{\mathbf{H}'} A_{\mathbf{H}} A_{\mathbf{H}'}^* e^{2\pi i \mathbf{H}' \cdot \mathbf{u}} \int_V e^{2\pi i (\mathbf{H}-\mathbf{H}') \cdot \mathbf{r}} dV(\mathbf{r}). \quad (11)$$

The integral on the right is always zero, except when $\mathbf{H} = \mathbf{H}'$, in which case the integral is equal to V . Using (8) we may thus write for (11):

$$\int_V \varrho(\mathbf{r}) \varrho(\mathbf{r}-\mathbf{u}) dV(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{H}} I_{\mathbf{H}} e^{2\pi i \mathbf{H} \cdot \mathbf{u}}. \dots \quad (12)$$

The right-hand side represents a Fourier summation, performed with intensities instead of amplitudes; compare (12) with (7). The significance of this Patterson function, as it is called, appears from the left-hand side: it is a volume integral of the product of two electron densities at places lying at vector distances \mathbf{u} apart. This implies that the integral in question will assume high values when \mathbf{u} is a distance between peaks in electron density, i.e. when \mathbf{u} is an interatomic distance really present in the crystal. By performing a Fourier synthesis, using the intensities of the reflections as Fourier coefficients, we thus obtain a function whose maxima provide indications of the atomic distances.

The clues to interatomic distances obtained by the Patterson method are sometimes so clear that they reveal the whole structure; sometimes they throw light only on certain details of the structure. Much depends on the ingenuity of the researcher's interpretation of the Patterson synthesis. However this may be, information is always present, which can help in the construction of the model.

We should mention finally that other methods exist which lead directly, i.e. not via models, to the phase constants required. One such is the "method of inequalities", which is based on the fact that the electron density in a crystal can nowhere be negative, and another is the "statistical method", where use is made of knowledge concerning the constituent

atoms of the crystal. Particulars of these methods will be found in the treatise by Bouman, mentioned in footnote⁴⁾.

These, then, are the broad lines of structure determination. We shall now say something about the "strategy".

We have seen that each reflection corresponds to a density component, providing information on its amplitude, wavelength and direction, and that to determine the phases a provisional model is devised. Now every incorrect model is shown to be wrong only after a timeconsuming procedure. The less probable models should therefore be avoided as far as possible. The time can better be used to look for a more reasonable model; again, the search for such a "good" model should not take too long. An important aspect of structure analysis is thus the art of finding, with the available data, the shortest way to this "good" model. Since the data involved are different in every case, one should always be prepared for other possibilities and difficulties, and be ready to try one way after another. The question often arises later whether one should persist with a choice once made, when the work based on the model adopted does not lead directly to results. In this way, experimenting and doubting, the crystallographer may be busy for months before he sees the first signs of success, indicating that he may at last start on the actual determination of the structure. This element of uncertainty imparts to structure determinations the interest and excitement of adventure.

Examples

Having sketched the broad outlines of the X-ray analysis of crystal structure, we shall now give a more detailed description of three cases, the first two being quite simple and making only partial use of the above procedures, and the third relating to a complicated structure, which necessitated a broader approach.

Analysis of the compound $Al_{0.89}Mn_{1.11}$

It was known as early as 1900 that various alloys of manganese are ferromagnetic, even though manganese itself is not ferromagnetic (Heusler alloys⁸⁾). The compound $Al_{0.89}Mn_{1.11}$ having turned

out to be a material suitable for permanent magnets, a study was made of the relation between the magnetic saturation and the structure of this substance.

The analysis of this structure went very smoothly. The Debye-Scherrer diagram showed the (tetragonal) unit cell to be very small, which indicated that there were not many different possibilities as to the atomic positions. It further appeared from the density and the chemical formula that the unit cell could in fact contain no more than two atoms. Finally, the symmetry found left no other choice but to place the atoms at the corners and at the centre.

It is very rare that the atomic sites can be determined with such accuracy at such an early stage of an analysis. Usually, it may be recalled, only the approximate positions of the atoms would be known at this stage, and on the basis of this rough model one would try by Fourier syntheses to obtain a more accurate picture of the structure. In this case that was not necessary because the symmetry had already pointed the way to the precise positions for the atoms.

This would have concluded the structural analysis had it not been that the non-stoichiometric formula $Al_{0.89}Mn_{1.11}$ showed that the aluminium and manganese atoms could not possibly occur in a completely regular arrangement in the crystal. If, for example, all aluminium atoms were situated at the corners and all manganese atoms in the centres, the formula would be $AlMn$ exactly. This suggested that the atoms were not so strongly bound to their own position in the cell, in other words that aluminium atoms might well occupy some of the sites of manganese atoms, and vice versa (the latter case prevailing).

In such cases we want to know the ratios in which the various kinds of atoms occur at the various sites. This information, too, can be deduced from the intensities of the reflections. If in this case we calculate the reflection intensities on the basis of a unit cell that is everywhere identical — containing for example one manganese atom at site A (corner) and one aluminium atom at site B (centre) — the results obtained will differ from the observed values. The effect on the intensities of the random atomic distribution described is as if each lattice site contained not one single particular atom but fractions of different atoms. These fractions on one site must, of course, together add up to one, and if we add the fractions of the same atoms in their different positions, the sum must obviously be in agreement with the chemical formula. In our case we can therefore write:

⁸⁾ F. Heusler, Über Manganbronze und über die Synthese magnetisierbarer Legierungen aus unmagnetischen Metallen, Z. angew. Chem. 17, 260-264, 1904.

	Site A	Site B
Aluminium	r	$0.89 - r$
Manganese	$1 - r$	$0.11 + r$

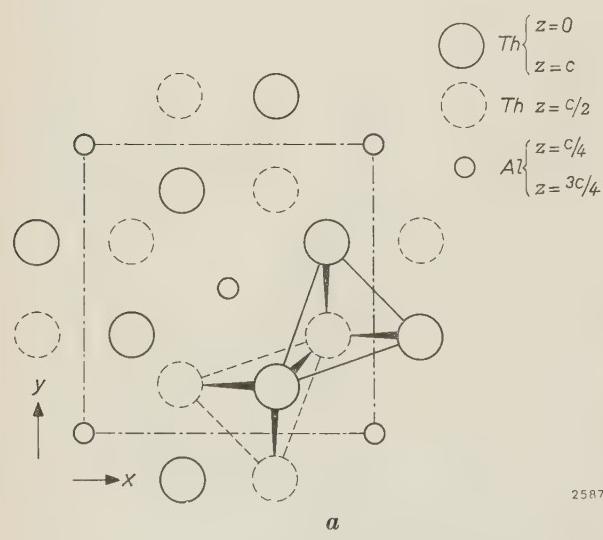
where r is the fraction of the number of A sites occupied by aluminium atoms.

We now calculate the reflection intensities for various values of r , and find the value of r that best agrees with the observed intensities. In our case this value was 0.03.

Finally, there was the crucial question whether this structure indeed corresponded to the magnetic properties of the substance. It was found, for example, that the magnetic saturation decreased when the substance was subjected to mechanical deformation. An X-ray investigation of this phenomenon revealed that r increases under these conditions to 0.13.

This meant that the drop in magnetic saturation must be accompanied by an increase in the number of manganese atoms on the "wrong" B sites (only the manganese atoms are considered, since they alone possess a magnetic moment). The conclusion was drawn that the magnetic moment of the manganese atoms on the B sites must be opposite in orientation to that of the manganese atoms on the A sites.

We shall not go further into this subject, but it may be mentioned that neutron diffraction analysis has confirmed this conclusion, making it possible to relate the magnetic saturation quantitatively with the above picture of the structure.



Analysis of the compound Th_2Al

In the course of research on thorium-aluminium compounds, which are important for their gettering properties, the structure of Th_2Al was analysed⁹⁾. This is an example of a fairly simple "trial and error" analysis, calling for no Fourier syntheses.

In this case the unit cell (again with tetragonal symmetry) contained four aluminium and eight thorium atoms. On grounds of crystal symmetry, the aluminium atoms could immediately be assigned to special positions in the cell, as was done for all atoms in the first example. It was not immediately clear where the eight thorium atoms should be located. The task of determining the 24 coordinates of these atoms was fortunately simplified by the fact that the symmetry severely limited the number of possibilities, inasmuch as a choice of one coordinate established the other 23.

The procedure thereafter was as follows. Various values of this coordinate were chosen, the reflection intensities were calculated with the models produced and the results were compared with the observed intensities. In this way the structure was determined by trial and error.

Here, too, the structure found threw light on to the other properties of the substance. The structure contains fairly large tetrahedral interstices (see fig. 8a and b). These interstices may be supposed to

⁹⁾ See also P. B. Braun and J. H. N. van Vught, Acta cryst. 8, 246, 1955.

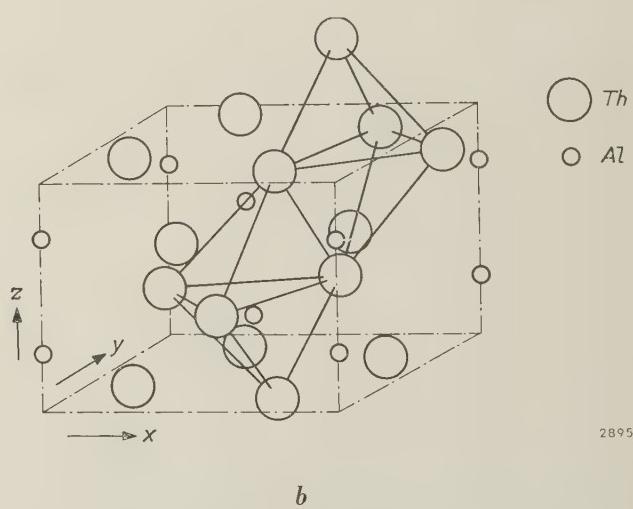


Fig. 8. The unit crystal cell of Th_2Al , a substance with gettering properties. The thorium atoms form tetrahedrons, a number of which are shown. Absorbed gas atoms or ions take up positions in the interstices of these tetrahedrons. a) Projected positions of the atoms looking along the c axis (5.86 \AA). The levels of the various atoms (z) are given in fractions of the c axis. The x and y directions are those of the a and b axes, respectively (both 7.62 \AA). b) Perspective view of unit cell, showing atomic positions. The x , y and z directions are again the a , b and c axes.

have some connection with the gettering properties of the substance. This supposition was later confirmed by nuclear magnetic resonance measurements¹⁰⁾ and by neutron diffraction analysis, from which it was possible to demonstrate the presence of gas atoms or ions in these interstices.

Analysis of the compound $\text{Y}-\text{Ba}_2\text{Zn}_2\text{Fe}_{12}^{\text{III}}\text{O}_{22}$

As our last example we shall take the analysis of a highly complex structure. $\text{Y}-\text{Ba}_2\text{Zn}_2\text{Fe}_{12}^{\text{III}}\text{O}_{22}$ is a substance from the class of ceramic magnetic materials which have been given the collective name "ferroxplana"¹¹⁾. These materials, whose properties resemble those of the familiar ferroxcube materials, can be used up to even higher frequencies than the latter (>100 Mc/s).

The fact that the unit cell in this case contains 114 atoms, viz. 6 barium, 36 iron, 6 zinc and 66 oxygen atoms, makes it evident that this analysis was a great deal more difficult than the two described above. The following outline of the analysis will give a good general picture of how such a complicated structure can be tackled.

One could not start on such a structure if it were not for the knowledge gained of simpler related structures. In this case, for example, we assumed that parts of the structure must resemble the structure of spinel¹²⁾.

With the aid of Debye-Scherrer patterns, Laue patterns and rotation photographs (the first obtained with an X-ray diffractometer) it was found that the crystal is hexagonal and that the unit cell has one very long edge (*c* axis, 43.6 Å) and two short edges (*a* and *b* axes, 5.9 Å).

Our first object was to find the position of the heaviest atoms, barium, iron and zinc. These contribute most to the scattering and therefore dominate the intensity pattern of the reflections. For this reason the position of such atoms must be accurately known. Only then is there any sense in looking for the sites of the lighter atoms. At this stage, however, no distinction could be made between iron and zinc atoms, the difference in their scattering power being relatively small.

It was assumed, from analogies and other indications, that the structure concerned would be one of layers perpendicular to the hexagonal axis, con-

sisting mainly of closely packed oxygen and barium atoms, most of the other atoms being in the interstices. First of all we tried to establish in which layers the barium atoms occurred. To do this it was sufficient to investigate a one-dimensional projection of the unit cell perpendicular to the hexagonal axis. The projection of the structure on to a plane was determined at a later stage, and finally the three-dimensional structure.

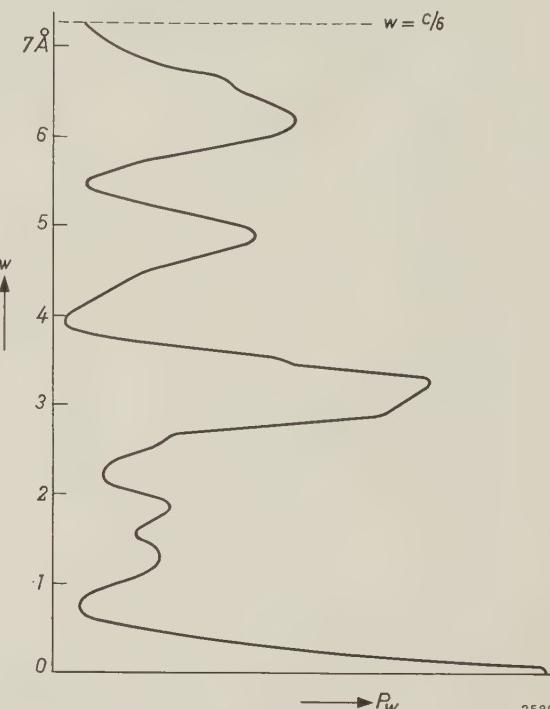


Fig. 9. Patterson plot of a one-dimensional Patterson synthesis $P(w)$ carried out with reflections from $\text{Y}-\text{Ba}_2\text{Zn}_2\text{Fe}_{12}^{\text{III}}\text{O}_{22}$. The *w* direction is that of the *c* axis. The various peaks provide information on the projections of the interatomic distances on the *c* axis. The peak at 3.2 Å indicated a projected spacing of that amount between barium atoms. This indication made it possible to build the first, provisional model.

A one-dimensional Patterson synthesis provided information on the projected distances between the atoms along the hexagonal axis (fig. 9). The large peak at 3.2 Å was taken as an indication of the projected spacing between the heavy barium atoms, || to the *c* axis. Knowledge of this spacing (however inaccurate due to the width of the peak) decided the choice between a number of models already drawn up on the basis of analogies. With the model thus obtained we set out to calculate the phases of the reflections.

This model represented, of course, a very rough structure, devoid of detail: all the light atoms and even part of the iron or zinc atoms had been disregarded. What is more, it was still only a projection. Using this incomplete model we determined the phases of 15 reflections and then carried out a Fourier synthesis. The resultant picture contained,

¹⁰⁾ See D. J. Kroon, Nuclear magnetic resonance, Philips tech. Rev. **21**, 286-299, 1959/60 (No. 10), in particular page 297.

¹¹⁾ See P. B. Braun, The crystal structures of a new group of ferromagnetic compounds, Philips Res. Repts. **12**, 491-548, 1957, and G. H. Jonker, H. P. J. Wijn and P. B. Braun, Philips tech. Rev. **18**, 145-154, 1956/57.

¹²⁾ Cf. E. J. W. Verwey, P. W. Haaijman and E. L. Heilmann, Philips tech. Rev. **9**, 185, 1947/48.

as we had hoped, numerous indications regarding the possible situations of the other atoms. This enabled us to venture on to new models, which were successively supplemented with more of the iron atoms still unused. When devising these models we made repeated checks with the Patterson synthesis (as we did on several occasions later) to ensure that the models chosen were not in conflict with the Patterson plot.

The model containing all the iron atoms was a good step forward, but still not sufficient to allow

us to distinguish between the iron and the zinc atoms and to find sites for the light oxygen atoms. We had the impression that this was because one of the iron atoms was not yet on its proper site. By analogy with the structures of other similar materials, we had allocated it a position in the same plane as the barium atoms. Various other positions were tried, at first with little success. Finally it was found that the answer was to place this iron atom just outside the plane containing the barium atoms. After this it was a fairly straightforward procedure to complete

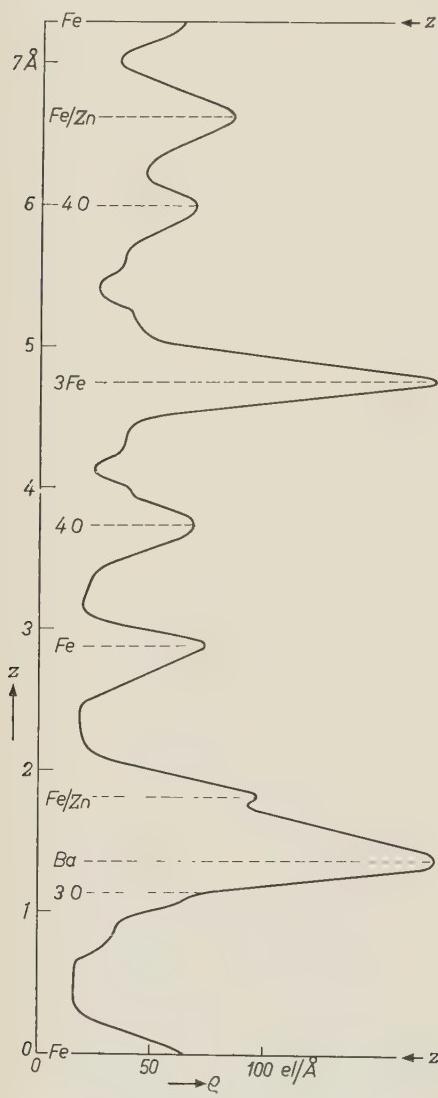


Fig. 10

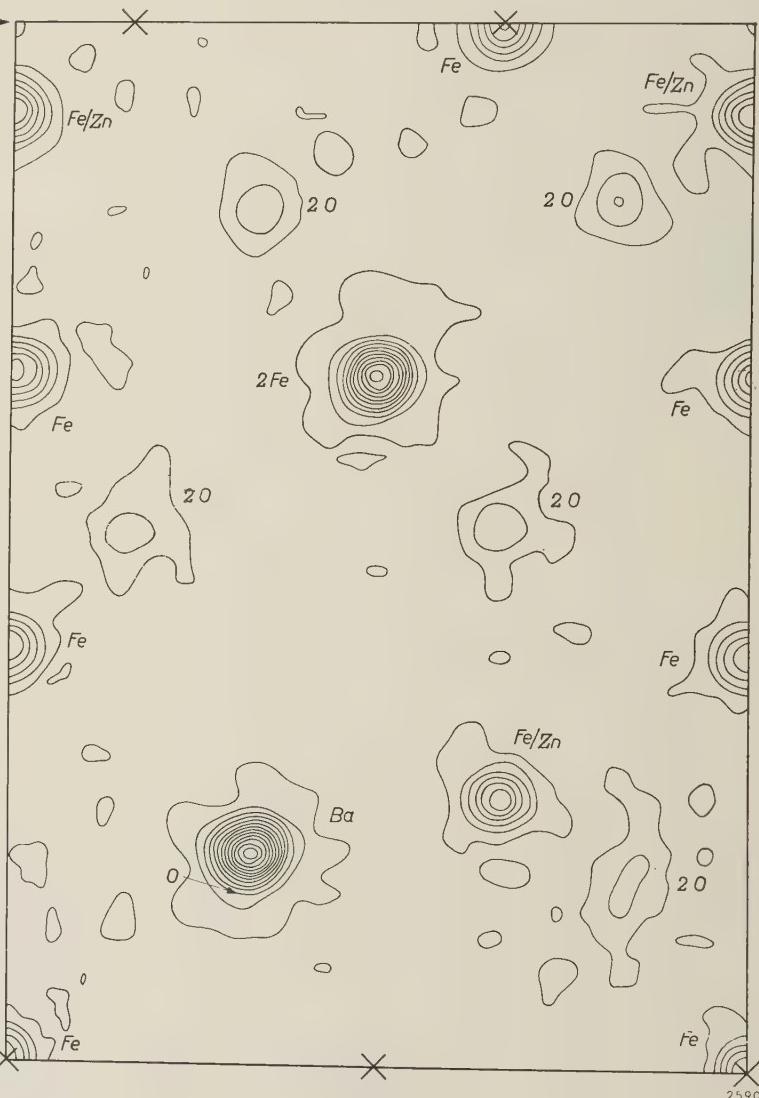


Fig. 11

Fig. 10. Result of a one-dimensional Fourier synthesis, obtained from the reflections from $\text{Y}-\text{Ba}_2\text{Zn}_2\text{Fe}_{12}\text{III}\text{O}_{22}$. The curve represents the variation of electron density (in electrons per ångström) in one sixth of the unit cell projected on to the c axis (z direction). The positions of the various projected atoms are denoted by their chemical symbols. This concluded the first stage of the analysis, in which only the projection on the c axis was considered. In order to reach this result, however, it was necessary to make many trials to determine the location of the iron (or zinc) atom, the peak of which is just perceptible beside the large peak for the barium atom. The next stage in the investigation is illustrated by fig. 11.

Fig. 11. Map of a two-dimensional Fourier synthesis, obtained from the reflections from $\text{Y}-\text{Ba}_2\text{Zn}_2\text{Fe}_{12}\text{III}\text{O}_{22}$. The content of one sixth of the unit cell is projected on to a plane. The contours shown represent lines of equal electron density. The "peaks" indicate the presence of atoms, denoted by their chemical symbols. The small peaks in the map are due to unavoidable experimental errors. The crosses are centres of symmetry. The z direction represents the c axis. The relation of the map to the projection of fig. 10 is seen directly simply by projecting the atoms on to the c axis. (It is only necessary to determine the one-sixth of the unit cell shown, because the rest of the unit cell can be obtained by inversion operations with respect to the centres of symmetry.)

the structure determination and extend it into two and three dimensions, thereby localizing the zinc and oxygen atoms. It is interesting to note that the zinc atoms and some of the iron atoms were found to compete for the same sites (exactly as in the case

of the manganese and aluminium atoms in $\text{Al}_{0.89}\text{Mn}_{1.11}$). In this way, then, the final structure was arrived at step by step. Figures 10, 11 and 12 mark the three most important steps of the determination.

The result of a one-dimensional Fourier synthesis (projection on the c axis), carried out after all the required phases had been calculated, is shown in fig. 10. Note the small peak beside the large peak of the barium atom: this is the iron atom that caused so much trouble. A two-dimensional Fourier synthesis — obtained analogously with the aid of two-dimensional models and the phases calculated on the basis of these models — gives the electron density of the structure projected on to a plane (fig. 11). (Projection of the peaks in this figure on to the c axis again produces the one-dimensional projection of fig. 10.) Finally, fig. 12 illustrates the three-dimensional structure found. Looking along the a axis we can recognize, between $z = 0$ and $z = c/6$, the projection given in fig. 11.

Attention may be called to some of the more important features of the structure thus determined. The iron atoms are located in octahedral interstices (atom C in fig. 12) or in tetrahedral interstices (atom B in fig. 12). Some of the tetrahedral

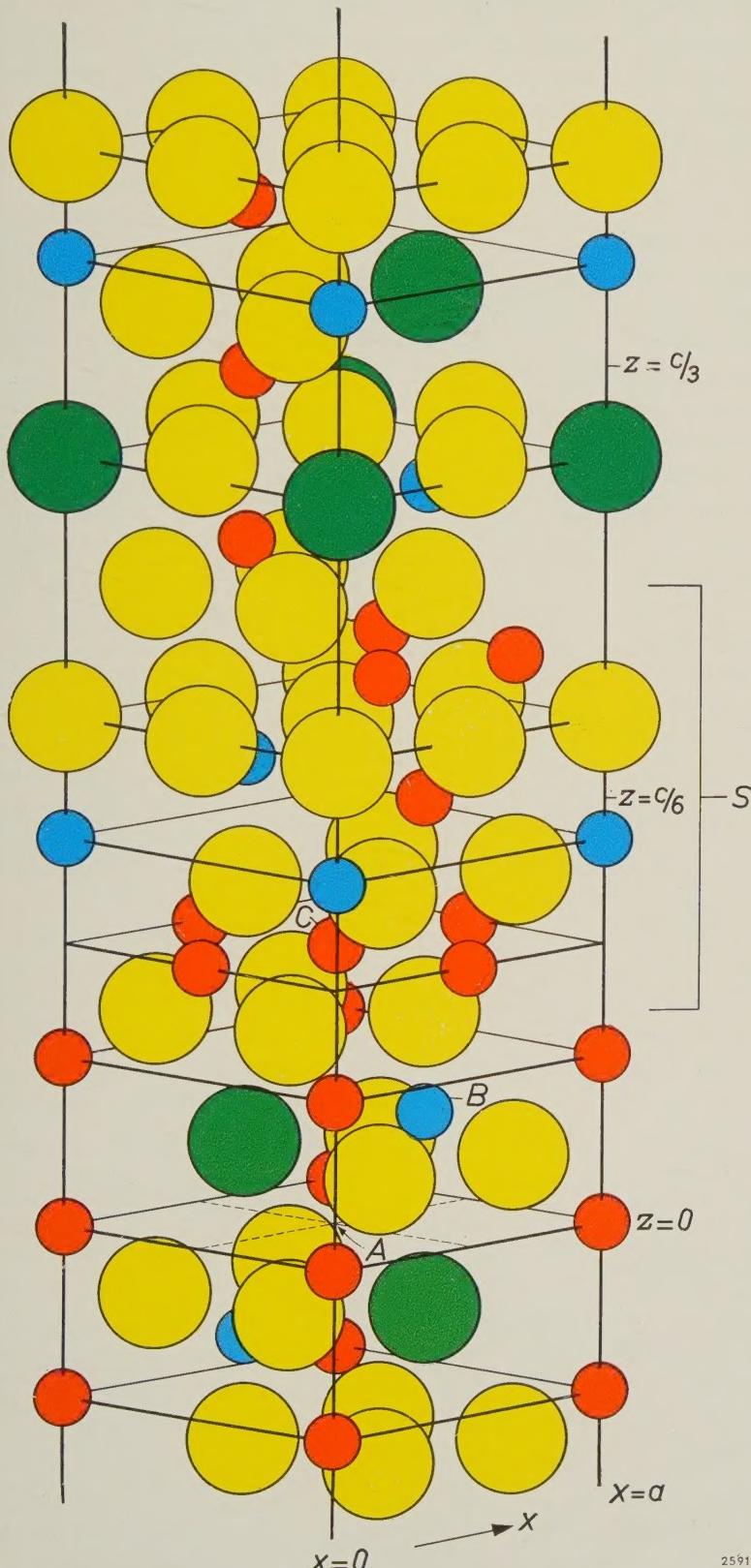


Fig. 12

Fig. 12. Three-dimensional model of part of the unit cell of $\text{Y}-\text{Ba}_2\text{Zn}_2\text{Fe}_{12}^{\text{III}}\text{O}_{22}$. The coloured spheres indicate the positions of the various atoms: green = barium atoms, red = iron atoms in octahedral interstices, blue = iron or zinc atoms in tetrahedral interstices, yellow = oxygen atoms. The long c axis runs in the vertical (z) direction. A projection of the cell between $z = 0$ and $z = c/6$ looking along the a axis (x direction) produces the map of fig. 11. In this three-dimensional representation it is possible to distinguish layers consisting of close-packed oxygen and barium atoms. The iron or zinc atoms are contained in certain of the oxygen interstices. Other important features of the structure are the presence of centres of symmetry (shown only at A , but others shown by \times in fig. 11) and of three 3-fold axes parallel with the c axis (not indicated). S denotes a block possessing the same structure as spinel. The fact that the interstices may be surrounded by either four or six oxygen atoms (tetrahedral or octahedral interstices) is clearly to be seen at B and C , respectively. B is also the iron (or zinc) atom which, as mentioned in the text, caused such trouble.

interstices are occupied by zinc atoms instead of iron atoms. Further, at either side of centres of symmetry (such as *A* in fig. 12) barium atoms are found which are relatively close together and are located in adjacent layers. (This had already been established, it may be recalled, from the result of the one-dimensional Patterson synthesis in fig. 9.) Owing to the proximity of these "large" atoms, the iron atom (or zinc atom) *B* is "pushed aside" slightly, so that this atom lies outside the plane, perpendicular to the *c* axis, through the layer of barium atoms (this, too, had emerged at an earlier stage, see fig. 10). ¹³⁾

The knowledge gained of the structure has been used for estimating the magnetic interactions between the various iron atoms, as described in this journal some years ago ¹¹⁾. It was found that all iron atoms on tetrahedral sites have parallel oriented magnetic moments, while those of the iron atoms on octahedral sites are for the most part anti-parallel. Such a distribution of the magnetic orientations explains the low magnetic saturation of this substance.

Another investigation concerned the property which this substance shares in common with other materials of the "ferroxplana" group, namely the existence of a preferred plane of magnetization. In this plane the magnetization is fairly free to rotate, enabling these substances to be used as a "soft"

¹³⁾ A full description is not possible in the compass of this article. No mention is made, for example, of the method of least squares, of difference Fourier synthesis, etc. A full description will be found in the article in Philips Research Reports quoted under ¹¹⁾.

magnetic material in rapidly alternating magnetic fields. The magnetic moments are tightly bound to the plane perpendicular to the *c* axis. Various calculations have been made to explain this effect, but they cannot be dealt with here ¹⁴⁾. It will suffice to remark that, as illustrated by the above "case histories", structure determinations can evidently lead to a better understanding of the properties of the substance investigated.

¹⁴⁾ See § 39 in "Ferrites" by J. Smit and H. P. J. Wijn, Philips Technical Library, 1959.

Summary. The principles and some applications of the X-ray determination of crystal structures are discussed. The electron distribution in a crystal can be described as the sum of a number of Fourier components, here termed "density waves". The diffraction by a crystal can be regarded as due to the scattering caused by these density waves. When a density wave is irradiated by X-rays, a perceptible reflection occurs only at a particular angle of incidence. The resultant reflections make it possible to determine the amplitude, wavelength and direction of these density waves. It then remains to determine the phase of the waves in order to calculate the electron distribution in a crystal by means of Fourier synthesis. The phases are found by successive approximation, on the basis of provisional models of the structure. Various means of arriving at such models are discussed, including the Patterson projection, which provides information on interatomic distances. Finally, various structure analyses carried out in the Philips Research Laboratories are described. Analysis of the structure of the compound $\text{Al}_{0.89}\text{Mn}_{1.11}$ made it possible to infer that the manganese atoms are present in two non-equivalent sites in the unit cell and that their magnetic moments at these sites are oppositely oriented. The structure of Th_2Al was found to contain fairly large interstices, which explained the gettering properties of this substance. The structure of $\text{Y}-\text{Ba}_2\text{Zn}_5\text{Fe}_{12}^{\text{III}}\text{O}_{22}$, a ceramic magnetic material (ferroxplana), proved to be exceptionally complicated (114 atoms in the unit cell). The successful analysis of its structure contributed to a deeper understanding of the magnetic properties of this substance.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS BY THE STAFF OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to Philips Electrical Ltd., Century House, Shaftesbury Avenue, London W.C.2, where a limited number of reprints are available for distribution.

- 2767:** M. Koedam: Sputtering of copper single crystals bombarded with monoenergetic ions of low energy (50-350 eV) (*Physica* **25**, 742-744, 1959, No. 8).

Experiments on the sputtering of copper atoms ejected from Cu single-crystal surfaces of various orientations bombarded with rare-gas ions at normal incidence show an anisotropy in the directional distribution. The ejection patterns in the case of a bombardment of Cu (111) and (110) surfaces are given.

The number of copper atoms ejected per incident ion in the normal (110) direction is given as a function of the ion energy for a copper (110) surface bombarded with rare-gas ions at perpendicular incidence. The number varies from 0.01 at 75 eV to about 0.04 at 350 eV, depending on the bombarding ion and its energy. For comparison the sputtering yield of silver is given as a function of the ion energy. The silver surface was bombarded with normally incident He^+ , Ne^+ , Ar^+ and Kr^+ ions, with energies ranging from 50 to 250 eV. The yield varies from 0.1 at 50 eV to about 1.5 at 250 eV ion energy.

- 2768:** A. Venema: The measurement of pump speed (*Le Vide* **14**, 113-120, 1959, No. 81; in French and in English).

Discussion of the various factors, in particular the definition and measurement of pressure, involved in the measurement of pump speeds at low pressures. A number of conclusions are reached concerning the location and nature of the pressure gauge to be used in order to obtain meaningful results.

- 2769:** K. H. Hanewald: Analysis of fat soluble vitamins, IV. Discussion of the chemical routine determination of vitamin D (*Rec. Trav. chim. Pays-Bas* **78**, 604-621, 1959, No. 8).

Details of the already published procedure for the chemical routine determination of vitamin D have been investigated, especially the colorimetric determination with antimony trichloride and the elimination of tachysterol by addition of maleic anhydride.

- 2770:** P. Westerhof and J. A. Keverling Buisman: Investigations on sterols, XII. The conversion of dihydrovitamin $D_2\text{-I}$ and $D_2\text{-II}$ into

dihydrotachysterol₂ (*Rec. Trav. chim. Pays-Bas* **78**, 659-662, 1959, No. 8).

The results are presented of investigation on preparing dihydrotachysterol₂ from dihydrovitamins $D_2\text{-I}$ and $D_2\text{-II}$.

- 2771:** E. J. W. Verwey: Synthetische keramiek (*Chem. Weekblad* **55**, 553-556, 1959, No. 41). (Synthetic ceramics; in Dutch.)

In the last decennia, especially in the last 10 years, there has been a rapid development of a new branch of ceramics, called here "synthetic ceramics", characterized by the use of synthetic starting materials and by the circumstance that the final composition corresponds to compound materials synthesized purposely in view of specific physical properties. These materials are mainly applied in high-frequency electrical technique. They comprise insulators, semiconductors and magnetic materials. A survey is given of this branch of ceramics. A number of materials developed in the Philips laboratories are discussed in some detail.

- 2772:** J. te Winkel: Drift transistor (*Electronic and Radio Engr.* **36**, 280-288, 1959, No. 8).

Equivalent circuits for a drift transistor are developed starting from a set of parameters derived from the physical principles underlying the device. It is shown what approximations are possible if limited frequency ranges or large values of the drift field are considered. Further simplifications are obtained from the introduction of suitably chosen frequency parameters. The resulting equivalent circuits appear to be simply related to those commonly used for a normal alloy transistor. The form is the same and the values of the circuit elements can be found by means of a number of multiplying factors that depend on the drift field only. These are given in graphical form.

- 2773:** C. Jouwersma: Die Diffusion von Wasser in Kunststoffe (*Chem. Ing. Technik* **31**, 652-658, 1959, No. 10). (The diffusion of water in plastics; in German.)

The application of classical diffusion theory to the water uptake of plastics is described. In the cases investigated experimentally, it was found that Fick's law was obeyed. In such cases, the

variables — viz. form and size of specimen, nature of surrounding medium, time and temperature, and solubility and diffusion coefficients of the material — can be separated. This enables one to make exact predictions concerning the water uptake of specimens of any form and size provided the characteristic constants of the plastic material in question are known.

2774: P. Penning and G. de Wind: Plastic creep of germanium single crystals in bending (*Physica* **25**, 765-774, 1959, No. 9).

The vertical displacement of the centre of a bar supported on two knife-edges and loaded in the centre, caused by plastic flow, is measured as a function of time. First the creep rate increases gradually and then becomes constant. The parameters describing the behaviour in these two regions have been determined as a function of the applied stress, for crystals of low and high dislocation densities and for crystals doped with oxygen.

2775: H. G. van Bueren: Theory of creep of germanium crystals (*Physica* **25**, 775-791, 1959, No. 9).

To explain the shape of the creep curves of germanium single crystals loaded in tension and in bending, a simple kinetic model is proposed, in which the dislocations are generated by (surface) sources and move with a uniform velocity over their glide planes. In this model a quantitative interpretation of the parameters of the creep curve in terms of the velocity of the dislocations, the incubation time of the sources and the density of the sources is possible. From the observations at "high" stress levels the velocity of dislocations in the germanium lattice can be determined; from those at "low" stress levels the rate of generation of the dislocations from the sources. The observed stress and temperature dependence of the creep process leads to similar dependences of incubation time and velocity. These dependences are used to form a quantitative theory of dislocation production and motion in the germanium lattice. This theory is shown to reflect semi-quantitatively various observed peculiarities of the creep phenomenon. The influence of other dislocations and of oxygen as an impurity on the elementary creep process can now also be qualitatively understood.

2776: J. S. C. Wessels: Dinitrophenol as a catalyst of photosynthetic phosphorylation (*Biochim. biophys. Acta* **36**, 264-265, 1959, No. 1).

The insensitivity of photosynthetic phosphorylation to dinitrophenol shows that the mechanisms of photosynthetic and respiratory generation of

ATP may be quite different. The present note reports that, surprisingly enough, DNP is able to catalyse the generation of ATP by illuminated chloroplasts. In this respect it is more active than FMN but less active than vitamin K₃.

2777: H. Koelmans, J. J. Engelsman and P. S. Admiraal: Low-temperature phase transitions in $\beta\text{-Ca}_3(\text{PO}_4)_2$ and related compounds (*Phys. Chem. Solids* **11**, 172-173, 1959, No. 1/2).

On measuring the temperature dependence of the luminescence of $\beta\text{-Ca}_3(\text{PO}_4)_2$ activated with divalent Sn, strong hysteresis effects were found. These effects are interpreted as being due to phase transitions occurring at about 20 °C and at about -40 °C. The presence of these two phase transitions has been confirmed calorimetrically. A number of modified Sr-orthophosphates, which have a structure closely related to that of $\beta\text{-Ca}_3(\text{PO}_4)_2$, show similar behaviour.

2778: J. Davidse: Transmission of colour television signals (*T. Ned. Radiogenootschap* **24**, 255-272, 1959, No. 5).

The paper discusses the transmission of colour-television signals according to the NTSC system. The choice of the chrominance signals, of their bandwidths and of the subcarrier frequency is discussed. The consequences of the method of gamma correction and of deviations from the constant-luminance principle are considered. The significance of the statistics of the chrominance signal is pointed out. This article has since been published in *I.R.E. Trans. on Broadcasting BC-6*, 3, Sept. 1960.

2779: A. G. van Doorn: Studio equipment for colour television (*T. Ned. Radiogenootschap* **24**, 237-253, 1959, No. 5).

This paper gives a broad survey of the equipment constructed in the Philips Research Laboratories for the generation of colour-television signals and the testing of the different pick-up systems. It describes three colour cameras, one using image orthicons as pick-up tubes, while in the other two experimental vidicons are used. Further the principle of the flying-spot scanner is described, as well as the colour-film camera. The special problems encountered in designing simultaneous pick-up systems and concerning colour-image registration, signal uniformity and gamma correction are discussed in more detail. In conclusion more is said about the different pick-up tubes and their use in colour-television cameras, their sensitivity, picture quality and overall performance.